Supply risk mitigation in a decentralized supply chain: pricing postponement or payment postponement?

Xin Geng
Miami Herbert Business School, University of Miami, xgeng@bus.miami.edu

Xiaomeng Guo, Guang Xiao
Logistics and Maritime Studies, Faculty of Business, Hong Kong Polytechnic University, xiaomeng.guo@polyu.edu.hk, guang.xiao@polyu.edu.hk

Nan Yang
Miami Herbert Business School, University of Miami, nyang@bus.miami.edu

Problem Definition: In a multi-stage model of a bilateral supply chain, we study two postponement strategies that the downstream retailer may adopt to mitigate the supply yield risk originated from the upstream production process. The retailer could either postpone the procurement payment until after the yield is realized and pay only for the delivered amount, or postpone the pricing decision to better utilize the available supply, or do both. Although both strategies have been separately studied in literature, there is little research on their combined effect and system-wide implications in a decentralized setting. *Methodology/Results*: Taking a game theoretic approach, we formulate a Stackelberg game and solve for the equilibrium in four scenarios respectively, in which the retailer uses different combination of the postponement strategies. There are three main findings. First, when the production cost is low and the yield loss is highly likely, the retailer never strictly benefits from either postponement strategy; with relatively high production cost, the retailer is more likely to adopt payment, rather than pricing, postponement. Second, we uncover a situation where postponing payment and postponing pricing are strategic complements for the retailer. That is, the use of one strategy may increase the benefit of using the other. Third, we identify conditions under which the postponement strategies can be Pareto optimal to the entire supply chain, making the firms' profits and the consumer surplus simultaneously higher. *Managerial Implications*: These results can be applied in many practical settings to provide guidance for firms to better design the procurement contract and properly use marketing instrument (pricing) to effectively mitigate supply risk and increase profit.

Key words: Supply risk, random production yield, postponement, pricing, price-only contract

1. Introduction

Outsourcing has become a trending and prevailing capacity planning strategy for many firms. In particular, firms outsource to cut costs, increase flexibility, enhance in-house efficiency, and ultimately build up competitive advantage in today's global market. Despite the apparent economic and strategic benefit, outsourcing can expose firms to various risks that arise externally; among others, supply risk stands out as one of the most notable and requires careful management. As documented by Gyorey et al. (2010), in responding to a McKinsey global survey, about two thirds of the participating executives reported that supply uncertainty was listed among the top three most significant risks to combat against. Commonly-seen across industries, the production yield risk is considered as a major source of supply risk. In agribusiness, the crop yield per farmed acre exhibits inherent variability and is usually subject to many unpredictable factors such as weather conditions, diseases infections, and seed mortality rate (Kouvelis 2015). In vaccine manufacturing, the production yield is subject to the survival rate of the viral strains due to the variable growing conditions (Chick et al. 2008). In the above examples, the uncertain yield has been a primary planning concern for every organization along the value chain.

Facing the random yield in the upstream production process, how should the downstream outsourcing firm manage and mitigate the risk? One practical way is to contract with the supplier to reduce the risk burden. Interestingly, achieving risk reduction does not always require sophisticated contract. Rather, a simple price-only contract would be sufficient if a certain payment scheme is stipulated. For example, in agriculture supply chains of crops such as cotton, soybeans, and potatoes, the downstream crop processors often contract with the upstream farmers in either "planted land acreage agreement", i.e., the payment is based on the total planted land acreage and the producers are required to deliver all available crops upon harvest, or "delivered output agreement", i.e., the payment is contingent upon final delivered outputs, rather than the planted land acreage (see page 54 in Scott 2003). Hence, the retailer can either pay for the ordered quantity (*pay-by-order* scheme) or pay for the actual delivered output after yield realization (*pay-by-delivery* scheme). In fact, the two payment schemes are quite popular in agricultural and livestock markets: Two types of contracts discussed in MacDonald et al. (2004), "production contract" and "marketing contract", correspond to pay-by-order scheme and pay-by-delivery scheme, respectively. In addition, both payment schemes are also widely adopted in hi-tech and biotech industries with significant supply yield risks; e.g., buyers in such industries may pay for either the startup production volume or the final delivered output (see, e.g., Tang and Kouvelis 2011, for detailed discussions). Note that, although the unit wholesale price is usually set by the supplier before yield is known, the retailer can have the decision right (or at least negotiation power) on which payment scheme to write in the contract upfront. Therefore, by using the pay-by-delivery scheme, the retailer can make the upstream supplier absorb part of the supply risk.

In addition to using contracts, the downstream retailer may adopt different pricing strategies to manage the yield risk. Endowed with pricing power, the retailer may either decide the price before production (*ex ante pricing*) or postpone the decision until after the yield realizes (*responsive* *pricing*). Both pricing schemes are widely observed in practice. Again, in agricultural and livestock business, according to Etienne et al. (2017), many agri-producers either pre-sell their processed products before production takes place (hence the input from the supplier is unknown yet), or price responsively after the yield is observed to make the best use of available supply. Similarly, in consumer electronic industry, Sinofsky (2019) documents that some companies may choose to pre-announce the prices of their products without making the products available immediately, whereas others may delay announcing such information until component supply risk is resolved and final products are ready for sell. Therefore, while an ex ante determined price can help maintain the stability of the output price, responsive pricing offers the retailer an opportunity to adjust the market price and alleviate the detrimental impact of the yield shortage on the end market.

While using different payment schemes deals with yield uncertainty from the supply side, using different pricing schemes addresses the same issue from the demand side. In practice, therefore, it is natural for firms to simultaneously consider the two means together as potential yield risk mitigation tools. One notable example in this regard is from the orange juice industry. Valencia oranges, a sweet orange cultivar perfect for making orange juice, are grown in South Florida and Brazil and sold to orange juice companies. In a general supply contract between the seller and the buyer of Valencia oranges[1], we can see the use of different payment schemes such as production contract (analogue to pay-by-order) and box contract (analogue to pay-by-delivery). On the demand side, the orange juice companies may raise the retail price when extreme weather and citrus diseases plague the orange crop and hurt the production yield[2]. More specifically, consider Tropicana, a unit of PepsiCo, Inc., sourcing Valencia oranges from its supplier Alico, Inc. and selling the processed orange juice to consumers. According to the purchase contract[3] and Alico's reporting[4], Tropicana buys the oranges on a pound solids basis, which is analogue to the pay-by-delivery payment scheme. On the other hand, when the crop size is unexpected disappointing due to bad weather such as a winter freeze, Tropicana usually chooses to directly raise the retail price (or indirectly downsize the juice jug) to cope with the upstream yield shortage[5].

Focusing on the downstream retailer in a two-tier supply chain setup, we study the above two risk mitigation strategies, both of which are related to the timing of events and emphasize the notion of

[1] Source: `www.sec.gov/Archives/edgar/data/1285785/000119312506170786/dex10iiy.htm`

[2] Source: `www.theguardian.com/science/2023/feb/13/florida-orange-juice-record-prices-weather-disease`

[3] Source: `contracts.justia.com/companies/alico-inc-6673/contract/118585/`

[4] Source: `ir.alicoinc.com/all-sec-filings/content/0000003545-19-000143/alco-93019x10k.htm`

[5] Source: `www.wsj.com/articles/SB10001424052748704655004575113593765269442`

postponement: A pay-by-delivery contract features the payment postponement, and the responsive pricing scheme postpones the retailer's pricing decision. Intuitively, both strategies could serve as effective levers for the retailer to combat supply yield uncertainty. However, in a decentralized supply chain setting, the combined effect and interplay of the two strategies are largely unclear. For one thing, due to channel conflicts, the upstream supplier's response to the retailer's strategies may inflict harm to the latter. For another, the joint effect of the two postponement strategies is not a simple addition, because of their interactions that ripple through the supply chain. Therefore, to ensure the maximum benefit of using these strategies, individually or combined, it is critical to understand their strategic implications in a decentralized setting.

The main purpose of this research is therefore to study the impact of the retailer's payment and pricing postponement strategies on its own profit achievement as well as the overall channel performance. Based on a Stackelberg game framework, we study four scenarios depending on the combination of the retailer's payment and pricing schemes and solve for the firms' optimal profits and the consumer welfare in equilibrium. Then, we attempt to answer the following research questions. (1) When should the retailer use what postponement strategy to improve its profit? This is the foremost choice to make for the retailer. (2) How would the use of one strategy influence the effect of using the other? This is to help the retailer understand the interplay between these two strategies. (3) What are the welfare implications of the postponement strategies on the overall channel performance? Knowing the answer to this question could help the retailer persuade other channel members to agree with its choices.

To answer the first research question, we examine the impact of the postponement strategies on the retailer's profit. Starting with the benchmark scenario where no postponement is present, we compare the retailer's optimal profit among the scenarios where the retailer deploys different combinations of the payment and pricing schemes. On the one hand, we find that, when the supplier's production cost is low and the likelihood of yield loss is relatively high, the benchmark scenario is already optimal; i.e., the retailer will not strictly benefit from either postponement strategy. If the production cost is high, however, the retailer's profit will be strictly improved if pricing postponement is used. On the other hand, we establish the retailer's preference order over various postponement strategy combinations. Specifically, with low production cost, the retailer is more likely to postpone payment rather than pricing; however, the retailer prefers the pricing postponement strategy when the supplier's production cost increases. Overall, we find that the benefit of pricing postponement for the retailer is growing as the production cost increases, because the contingent market price accompanied with low production cost would worsen the double marginalization problem. Furthermore, the retailer benefits less from postponing payment when the cost increases, because the supplier simply will not inflate production even under a pay-by-delivery contract if doing so is too costly.

In addition to the direct comparison among the combinations of strategies, we also investigate the interplay of the payment and pricing schemes to answer the second research question. Since they are both potential tools for the retailer to mitigate the supply yield risk, intuition suggests that the two postponement strategies may be substitutes. Nevertheless, our analysis uncovers the opposite: When the potential yield loss is not too severe and the production cost is in the medium range, payment postponement and pricing postponement are strategic complements. In such a situation, the use of one strategy can enhance the positive effect of the other for the retailer.

Finally, for the last research question, we scrutinize the system-wide implications of the retailer's postponement strategies. To be specific, we take the channel's perspective, and study the respective impacts of the payment and pricing schemes on firms' profits as well as the consumer surplus. We find that payment postponement can be a Pareto optimal choice for the entire supply chain, especially when the supplier's production cost is low. However, with relatively high cost, either the consumers (without pricing postponement) or the supplier (with pricing postponement), but not both, will be negatively affected by the payment postponement. As for the impact of the pricing postponement, we show that it can be positive for the entire channel, but the benefit is largely enhanced when accompanied by payment postponement. In fact, the retailer's pricing postponement strategy may hurt the supplier if payment is not postponed. Rather, once the retailer shifts the yield risk back to the upstream and lets the supplier make inflation decision, postponing the pricing decision will never hurt anyone in the supply chain. Under each payment scheme, pricing postponement is more likely to benefit the supply chain with relatively high production cost.

The rest of the paper is organized as follows. Section 2 surveys the extant literature and positions our work. Section 3 formulates and analyzes the model. We study the retailer's strategy selection problem in Section 4, and investigate the channel implications in Section 5. Section 6 presents numerical studies, Section 7 discusses three interesting extensions, and finally Section 8 concludes the paper. Proofs and additional results are all relegated to the appendices.

2. Literature Review

Since the primary focus of this paper is to identify effective mechanisms to cope with supply shortage, it is related to the literature on supply risk management in general. Depending on the nature of the risk, there are three commonly adopted supply risk models in previous research: Random capacity model is used when supply shortage is due to unexpected capacity deterioration (e.g., Ciarallo et al. 1994), binomial disruption models severe supply problems caused by disruptive events (e.g., Chen et al. 2001), and proportional random yield model captures the output shortage led by the inherent risk suffered by the production process (e.g., Yano and Lee 1995, Grosfeld-Nir and Gerchak 2004). Our paper uses the proportional random yield model because it best fits the motivating setting of agribusiness. This line of research is largely devoted to the design of operational strategies to effectively mitigate yield risk. For example, firms can inflate production and/or hold extra inventory to hedge against yield risk (see Henig and Gerchak 1990, Li and Zheng 2006, Kouvelis et al. 2018), diversify their supply base to enjoy the risk pooling effect and reduce supply output variability (see Anupindi and Akella 1993, Dada et al. 2007, Federgruen and Yang 2008, 2009, 2011, 2014, Hu and Kostamis 2015, Tang and Kouvelis 2011, Dong et al. 2016, 2022), and exert effort to improve the suppliers' production reliability (see Wang et al. 2009, Tang et al. 2014, Wang et al. 2014).

Both risk-mitigating strategies studied in this paper highlight the postponement of certain event, and thus our paper is related to the literature that examines the effect of payment/pricing postponement on firm's operational measures. In this regard, the pricing postponement strategy has been studied in various settings. First, previous research has considered variability from the demand side. Van Mieghem and Dada (1999) first study the impact of the pricing timing on the inventory decision under demand uncertainty. Chod and Rudi (2005) investigate how responsive pricing affects the usage of flexible resources under correlated demands. Granot and Yin (2008) study how order and price postponement affect the performance of a decentralized supply chain with a price-setting newsvendor. Second, in the presence of random yield, Tang and Yin (2007) and Kouvelis et al. (2021) illustrate the benefit of pricing postponement for risk neutral and risk averse firms, respectively. Li et al. (2017) and Dong et al. (2016) evaluate the strategic interaction between diversification and responsive pricing in mitigating either supply capacity risk or yield risk.

As for payment postponement, it has been widely studied in the supply chain finance literature that concerns trade credit related contracts. This stream primarily focuses on understanding the financing role of payment postponement and how it affects channel efficiency (see, e.g., Kouvelis and Zhao 2012, Chod 2017, Yang and Birge 2018, Chod et al. 2019, etc). In the context of supplier's product failure, Babich and Tang (2012) and Rui and Lai (2015) investigate how to use payment postponement, together with inspection, to mitigate supplier product adulteration risk. In addition, Tang et al. (2018) study how payment postponement interacts with other financing mechanisms to reduce financially constrained supplier's performance risk under asymmetric information. Different from the above works, we consider the setting where the supplier may deliver a random fraction of the order quantity, in contrast to the all-or-nothing delivery performance assumed in the above papers. As such, payment postponement could incentivize the supplier to voluntarily inflate the production quantity in order to improve the final delivery performance.

In summary, despite the vast volume of extant works on random yield and how firms cope with the risk by maintaining certain flexibility, we have not seen payment postponement and pricing postponement being studied together. However, in practice, the retailer facing random yield in the upstream production process is able to use both postponement strategies. Hence, results from prior works can only describe the optimal strategy for the retailer from one aspect. For example, responsive pricing may be shown to benefit the firm in some setting, but it may become harmful if a different payment scheme is used. Therefore, our foremost contribution is to investigate the strategic interaction between payment postponement and pricing postponement in mitigating supply yield risk in a decentralized supply chain, and to provide a more comprehensive strategy analysis for firms to manage yield risk in practice.

3. Model and Analysis

3.1. Basic Framework

Consider a supply chain that comprises one supplier and one retailer. The retailer orders the product from the supplier and sells to an end market of consumers. The supplier's production yield is subject to variability, and as a result, the planned production level will not always be fully realized. Let Y be the stochastic proportional yield factor with the following distribution:

$$Y = \begin{cases} 1 & \text{with probability} \quad \theta; \\ \alpha & \text{with probability} \quad 1-\theta. \end{cases}$$

Here, $\alpha \in (0,1)$ measures the severity of yield loss and $\theta \in (0,1)$ represents the likelihood of perfect yield. Let $\mu = \alpha(1-\theta)+\theta$ be the expected yield factor. One practical example captured by our yield model is the livestock and crop supply process, where the animals or plant seeds may catch infectious diseases with probability $1-\theta$ and the mortality rate is $1-\alpha$. Thus, our model serves as a sensible estimate when yield loss is related to defect rate or survival rate. The binary random yield factor, despite its simple form, is able to capture both the severity and likelihood of yield loss, allowing us to characterize the strategic interplay of postponement strategies with an insightful and tractable model. Finally, we assume that the production cost, c per unit, is incurred to the supplier for the entire planned production, regardless of the realized yield.

On the demand side, we assume that the end market is deterministic and price-driven. Specifically, given a retail price p, the demand for the product is $D(p) = a - bp$, where a is the potential market size and b is the price sensitivity. The linear functional form is widely adopted in the literature (e.g., Tang and Kouvelis 2011, Dong et al. 2022). In the main model, we assume deterministic demand to fully focus on the impact of the supply uncertainty. We relax this assumption and incorporate demand uncertainty in one of our extensions (see Section 7.2).

3.2. Postponement Strategies

As mentioned in Section 1, the retailer may use two postponement strategies as potential risk-mitigating tools. First, the timing of the procurement payment could be specified in a price-only contract. The *pay-by-order* scheme requires the retailer to complete the transaction when placing the order, based on the ordered quantity q. The *pay-by-delivery* scheme, on the other hand, postpones the payment until after the yield realizes and requires the retailer to pay only for the delivered quantity. Due to its contingency nature, the latter intuitively shifts the yield risk to the upstream supplier. Indeed, under a pay-by-order contract, the supplier's production level Q will be exactly the same as the ordered quantity q; however, under a pay-by-delivery contract, the supplier may have incentives to inflate the production. Hence, the delivered quantity, which is a random variable, is given by $\hat{q} = \min\{YQ, q\}$. Following the common practices in agribusiness (see Scott 2003, MacDonald et al. 2004), we assume that the retailer sets the payment scheme in the contract and the supplier sets the wholesale price.

Second, the retailer may use two pricing strategies when selling to the end market: Either committing to a retail price before the random yield realizes (referred to as *ex ante* pricing), or deciding a retail price at a later stage when the actual yield is known (referred to as *responsive* pricing). In the first case, since the retail price decided before yield realization may not be able to match supply with demand, the retailer has to absorb the supply risk by itself. In the second case, the retailer adopts responsive pricing and is thus able to transfer part of the supply risk to the downstream consumers by adjusting the retail price according to the available supply. Note that the retail price p in this setting is decided based on the available-for-sale amount, i.e., the delivered quantity $\hat{q}$, so that the retailer's total revenue, $p\min\{d(p), \hat{q}\}$, is maximized.

3.3. Problem Formulation

To examine the impact of the above postponement strategies, we look at four scenarios: (1) Scenario OA (pay-by-order and ex ante pricing); (2) Scenario OR (pay-by-order and responsive pricing); (3) Scenario DA (pay-by-delivery and ex ante pricing); (4) Scenario DR (pay-by-delivery and responsive pricing). In this paper, we will use these abbreviations as superscripts to differentiate the scenarios; moreover, see a list of notations in Appendix A for ease of reference. In each scenario, we model the interactions between the supplier and the retailer via a Stackelberg game. We assume a symmetric information setting, i.e., all the yield, cost, and demand information is common knowledge. This is a popular assumption in the related literature (e.g., Granot and Yin 2008, Babich and Tang 2012, Tang and Kouvelis 2011, Kouvelis et al. 2021) and will let us focus on answering the main research questions. Depending on the strategies adopted by the retailer, the sequence of events along the channel differs; see Table 1 for an illustration.

Table 1 The Sequence of Events in Each Scenario.

OA	OR	DA	DR
Supplier decides wholesale price w;			
Retailer orders q from supplier, pays for the ordered amount, and sets retail price p;	Retailer orders q from supplier and pays for the ordered amount;	Retailer orders q from supplier and sets retail price p;	Retailer orders q from supplier;
Supplier decides production level Q and produces;			
Yield realizes and supplier delivers $\hat{q}=\min\{YQ,q\}$ to retailer;			
Retailer sells to the end market.	Retailer sets retail price p and sells to the end market.	Retailer pays for the delivered amount and sells to the end market.	Retailer pays for the delivered amount, sets retail price p, and sells to the end market.

Given a certain scenario, we formulate the retailer's and the supplier's problems following the sequence of events shown in Table 1. We assume that both firms are risk neutral. Hence, the retailer chooses the order quantity and the retail price to maximize its expected profit, which is the revenue from the end market less the procurement cost. Specifically, the revenue from selling to the consumers is given by $r(p,q;Q)=p\min\{d(p),YQ,q\}$ and the procurement cost is simply a transfer payment T^k in scenario k. Under the price-only contract, $T^k=T^k(w,Q,q)$; specifically, $T^k=wq$ if $k\in\{OA,OR\}$ (pay-by-order), and $T^k=w\min\{YQ,q\}$ if $k\in\{DA,DR\}$ (pay-by-delivery). Moreover, the supplier will (sequentially) choose the wholesale price and the production level to maximize its expected profit, which is T^k-cQ.

We now formulate each firm's problem in a backward fashion. First, given the wholesale price

w and the retailer's order quantity q, the supplier decides production quantity $Q = Q(w,q)$ to maximize its expected profit:

$$\max_{Q} \mathbf{E}\left[T^k(w,Q,q)\right] - cQ, \quad \text{for } k \in \{OA, DA, OR, DR\}.$$

Next, given (w,Q), the retailer solves the following problem in each scenario:

$$\begin{cases} \max_q \mathbf{E}\left[\max_p r(p,q;Q(w,q)) - T^k(w,Q(w,q),q)\right] & \text{if } \ k \in \{OR, DR\}; \\ \max_{q,p} \mathbf{E}\left[r(p,q;Q(w,q)) - T^k(w,Q(w,q),q)\right] & \text{if } \ k \in \{OA, DA\}. \end{cases}$$

Lastly, anticipating the retailer's decisions $q(w)$ and the corresponding production $Q(w) := Q(w,q(w))$, the supplier optimizes its expected profit by setting the wholesale price w:

$$\max_{w} \mathbf{E}\left[T^k(w,Q(w),q(w))\right] - cQ(w), \quad \text{for } k \in \{OA, DA, OR, DR\}.$$

In addition to the firms' profits, we are also interested in the impact of the postponement strategies on the consumers. To that end, we first define the consumer surplus under uncertain supply. Let X be the available-for-sale amount and p be the market price. Then, define

$$\mathcal{S}(p,X) = \int_0^{\min\{d(p),X\}} (p(t) - p)dt,$$

where $p(d)$ is the inverse demand function. In above formula, we implicitly assume that when the supply shortage happens, the available inventory is allocated with priority to the consumers with higher willingness-to-pay, which generates the highest possible consumer surplus among other allocation rules (Cohen et al. 2018). In addition, by our assumption on the demand, there exists a unique d^* that maximizes the revenue function $p(d)d$. Hence, given X, the optimal ex post market price is decided by $p^*(X) = p(d^*)\mathbf{1}_{\{X \geq d^*\}} + p(X)\mathbf{1}_{\{X < d^*\}}$.

In our setting, the random variable $X = \hat{q} = \min\{YQ, q\}$, i.e., the delivered quantity. Therefore, after solving for the market price p^k, order quantity q^k, and production quantity Q^k in equilibrium, we may derive the expected equilibrium consumer surplus in each scenario:

$$\begin{cases} \mathbf{E}\left[\mathcal{S}(p^k, \hat{q}(q^k,Q^k))\right] & \text{if } \ k \in \{OA, DA\}; \\ \mathbf{E}\left[\mathcal{S}(p^*(\hat{q}(q^k,Q^k)), \hat{q}(q^k,Q^k))\right] & \text{if } \ k \in \{OR, DR\}. \end{cases} \tag{1}$$

We denote the equilibrium profits of the retailer and supplier and the corresponding consumer surplus in scenario k by π_R^k, π_S^k, and CS^k respectively. A direct observation from the problem formulation is that the equilibrium results in each scenario k, $(\pi_R^k,\ \pi_S^k,\ CS^k)$, depend on the production cost c, the yield distribution (α,θ), and the demand curve characteristics (a,b).

3.4. Equilibrium Results

We adopt the sub-game perfect equilibrium concept and use backward induction to solve the problem. The equilibrium results for each scenario are summarized below.

Lemma 1 **(Scenario OA Equilibrium)**. *The retailer's profit* π_R^{OA}*, the supplier's profit* π_S^{OA}*, and the consumer surplus* CS^{OA} *at equilibrium, respectively, are uniquely defined and have three continuous pieces on* $0 \leq c < \max\{0, c_1^{OA}\}$, $\max\{0, c_1^{OA}\} \leq c < \max\{0, c_2^{OA}\}$ *and* $\max\{0, c_2^{OA}\} \leq c < \frac{a}{b}\mu$*; the critical values* $c_1^{OA} < c_2^{OA} < \frac{a}{b}\mu$ *and the equilibrium results are specified in Appendix B.1. Moreover, the retailer inflates its order quantity only when* $0 \leq c < \max\{0, c_2^{OA}\}$*, and the supplier never inflates its production, i.e.,* $Q = q$.

As analytically characterized and graphically illustrated in Appendix B.1, there exists a partition of the unit square that consists of three regions, $\mathcal{H}_i$ $(i = 1, 2, 3)$, such that $0 < c_1^{OA} < c_2^{OA}$ if $(\alpha, \theta) \in \mathcal{H}_1$; $c_1^{OA} \leq 0 < c_2^{OA}$ if $(\alpha, \theta) \in \mathcal{H}_2$; and $c_2^{OA} \leq 0$ if $(\alpha, \theta) \in \mathcal{H}_3$. We will use this partition for our qualitative discussions later.

Lemma 2 **(Scenario OR Equilibrium)**. *The retailer's profit* π_R^{OR}*, the supplier's profit* π_S^{OR}*, and the consumer surplus* CS^{OR} *at equilibrium, respectively, are uniquely defined and have two continuous pieces on* $0 \leq c < \max\{0, c^{OR}\}$ *and* $\max\{0, c^{OR}\} \leq c < \frac{a}{b}\mu$*; the critical value* $c^{OR} < \frac{a}{b}\mu$ *and the equilibrium results are specified in Appendix B.2. Moreover, the retailer inflates its order quantity only when* $0 \leq c < \max\{0, c^{OR}\}$*, and the supplier never inflates its production, i.e.,* $Q = q$.

Lemma 2 implies that the threshold value c^{OR} can be non-positive. From the proof, we can see that either $\alpha \geq 1/2$ or $\theta \geq 1/2$ is sufficient to ensure that $c^{OR} \leq 0$. In this case, there is no order inflation from the retailer. Hence, when the likelihood of experiencing yield loss is low, or when the possible shortage from the random yield is small, using responsive pricing to mitigate the risk is already adequate. However, if the risk is both likely and severe and the production cost is sufficiently low, it is optimal for the retailer to inflate its order.

Lemma 3 **(Scenario DA Equilibrium)**. *The retailer's profit* π_R^{DA}*, the supplier's profit* π_S^{DA}*, and the consumer surplus* CS^{DA} *at equilibrium, respectively, are uniquely defined and have two continuous pieces on* $0 \leq c < c^{DA}$ *and* $c^{DA} \leq c < \frac{a}{b}\mu$*; the critical value* c^{DA} *and the equilibrium results are specified in Appendix B.3. Moreover, the supplier inflates the production level, i.e.,* $Q = q/\alpha$*, only when* $0 \leq c < c^{DA}$*, and the retailer never inflates its order.*

Lemma 3 confirms that, under a pay-by-delivery scheme, the supplier absorbs the yield risk by resorting to production inflation whenever profitable. Moreover, unlike the previous scenarios, in scenario DA, the equilibrium results as functions of c always have two pieces. This means that regardless of the yield distribution, inflation is always used to combat the yield risk when the production cost is low.

Lemma 4 **(Scenario DR Equilibrium)**. *Suppose $\alpha \geq 9/16$. The retailer's profit π_R^{DR}, the supplier's profit π_S^{DR}, and the consumer surplus CS^{DR} at equilibrium, respectively, are uniquely defined and have three continuous pieces on $0 \leq c < c_1^{DR}$, $c_1^{DR} \leq c < c_2^{DR}$, and $c_2^{DR} \leq c < \frac{a}{b}\mu$; the critical values c_i^{DR} ($i = 1, 2$) and the equilibrium results are specified in Appendix B.4. Moreover, the supplier inflates the production ($Q = q/\alpha$) when $0 \leq c < c_1^{DR}$, but does not inflate the production ($Q = q$) when $c_2^{DR} \leq c < \frac{a}{b}\mu$. When $c_1^{DR} \leq c < c_2^{DR}$, the supplier is indifferent to inflation, and we assume $Q = q/\alpha$ since the resulting consumer surplus is higher. The retailer never inflates its order.*

In Lemma 4, we assume $\alpha \geq 9/16$ to get clean analytical results. Indeed, in a sole sourcing setting such as ours, it is plausible to assume that the retailer only sources from a relatively reliable supplier. Note that the general case can be analyzed and the results are given in the proof in Appendix B.4.

4. Retailer's Strategy Selection

In this section, we take the retailer's perspective and investigate the selection of the postponement strategies. Our analysis will pivot on two aspects of the retailer's strategy selection. First, starting from scenario OA, which is assumed to be the *status quo* scenario, should the retailer choose to move to other scenarios (OR, DA, DR), and if yes, which one? Second, since the two strategies are both designed to mitigate the supply risk for the retailer, would using one postponement strategy enhance or diminish the effect of using the other?

4.1. Preference Between Various Postponement Strategies

Starting from scenario OA, we examine the retailer's preference between different scenarios. To obtain organized insights, we tackle this problem from two standpoints. First, we derive sufficient conditions under which the status quo scenario dominates all other scenarios. Thus, in this case, there is no need to consider other scenarios. Second, we assume that scenario OA is suboptimal. Then, we compare the other scenarios and find the best one for the retailer. Note that, if the retailer has the same profit in two scenarios, we assume that it prefers the scenario with less postponement; i.e., the adoption of any postponement strategy occurs only with strict dominance.

Should the retailer depart from scenario OA? The retailer will not even consider any postponement strategy if the status quo already leads to the highest profit. Otherwise, if the current profit is *strictly* dominated by the profit under any of scenarios OR, DA, or DR, the retailer should consider using at least one postponement strategy. The following proposition provides sufficient conditions in this regard.

PROPOSITION 1. *Suppose* $\alpha \geq 9/16$. *Then, there exists a* $0 < \tilde{C} < \frac{a}{b}\mu$ *such that*

(i) if $(\alpha, \theta) \in \mathcal{H}_1$, *then* $\pi_R^{OA} = \pi_R^{DA} = \pi_R^{DR} > \pi_R^{OR}$ *for* $0 < c < \min\{\tilde{C}, c_1^{OA}\}$;

(ii) if $(\alpha, \theta) \in \mathcal{H}_2$, *then* $\pi_R^{OA} > \pi_R^{DA} = \pi_R^{DR} > \pi_R^{OR}$ *for* $0 < c < \min\{\tilde{C}, c_2^{OA}\}$.

On the contrary, for any $0 < \alpha, \theta < 1$ *and* $\max\{0, c_2^{OA}\} < c < \frac{a}{b}\mu$, *OA is strictly dominated by at least one of the other three scenarios.*

Proposition 1 identifies two conditions under which the status quo scenario OA is already dominating. Under either condition, the production cost is relatively low and the parameter θ is not large (see the definition of $\mathcal{H}_i$, $i = 1, 2, 3$, in Appendix B.1). Moreover, in these cases, solely postponing pricing is always a strictly dominated strategy, and using pay-by-delivery contract to postpone payment results in the same retailer's profit regardless of the pricing schemes. This finding shows that, when the channel is relatively cost-efficient (i.e., small c) and the likelihood of having yield loss is high (i.e., small to medium θ), the retailer is better off by maintaining the status quo; any strategic postponement would result in unchanged or lower profit.

Let us understand the above result from the perspective of the supplier's reaction to the retailer's decision. Since the production cost is low, the retailer is likely to order more. If the retailer uses responsive pricing to take advantage of a more efficient (and perfectly matched) retail market, the supplier would also want to benefit from it. Thus, the responsive pricing will lead to a higher wholesale price and worsen the double marginalization problem. In fact, with relatively unreliable yield, the retailer will actually be hurt by moving to scenario OR. If the retailer chooses a pay-by-delivery contract to transfer the yield risk back to the supplier, this decision has two unfavorable consequences. First, the retailer loses the control in the inflation decision, which may be manipulated to benefit itself. Second, the supplier may charge a "premium" on the wholesale price to compensate for bearing the risk. Therefore, due to the unreliable yield, it is difficult for a pay-by-delivery contract to benefit the retailer. Moreover, note that in $\mathcal{H}_2$, θ is in the medium range, so the yield variance is large. As a result, the supplier may increase the wholesale price even more, resulting in the strict dominance described in the above proposition.

Lastly, Proposition 1 also gives a situation where the retailer will want to depart from the status quo scenario. Specifically, we find that responsive pricing makes the retailer strictly better when the supplier's production cost is relatively high. Therefore, for a less cost-efficient supply chain, it is better for the retailer to deploy at least one postponement strategy (responsive pricing). It is worth noting that, $c_2^{OA} < 0$ in $\mathcal{H}_3$, and therefore, by Proposition 1, scenario OA is strictly dominated for all feasible c. This means that the supplier's reliability has a considerable impact on the retailer's strategy selection. When the likelihood of supply shortage $(1-\theta)$ is small and thus the supplier is relatively reliable, the retailer would prefer leaving the status quo scenario even if the production cost is low.

Suppose that scenario OA is not optimal for the retailer. In this case, $\pi_R^{OA} < \max\{\pi_R^k; k = OR, DA, DR\}$. Then, to strictly improve profit, the retailer needs to move away from the current scenario by postponing either the procurement payment or the retail pricing, or both. Hence, we compare the retailer's equilibrium profit under scenario OR, DA and DR, respectively, to find out when to use what strategy. The supplier's response to the retailer's strategic postponement, which will affect the outcome of the vertical interaction in the decentralized supply chain, is largely driven by the yield uncertainty and the production cost. The next proposition characterizes the retailer's preference among the various postponement strategies.

Proposition 2. *Let $\alpha \geq 9/16$. Then the following statements hold.*

(i) $\pi_R^{OR} < \pi_R^{DR} = \pi_R^{DA}$ *for* $0 < c < \tilde{C}$.

(ii) $\pi_R^{DA} = \pi_R^{DR} < \pi_R^{OR}$ *for* $\tilde{C} < c < c_1^{DR}$.

(iii) $\pi_R^{DA} < \pi_R^{OR} < \pi_R^{DR}$ *for* $c_1^{DR} < c < c_2^{DR}$.

(iv) $\pi_R^{DA} < \pi_R^{DR} = \pi_R^{OR}$ *for* $c_2^{DR} < c < \frac{a}{b}\mu$.

As the production cost c increases, the preferred postponement strategy changes for the retailer. When $0 < c < \tilde{C}$ (threshold $\tilde{C}$ was introduced in Proposition 1), the optimal scenario is DA. That is, the retailer is better off using a pay-by-delivery contract to postpone procurement payment. The pricing postponement, on the other hand, should not be considered. This shows that in a relatively cost-efficient supply chain, the retailer would rather transfer the yield risk to the supplier than to the end-market consumers. Indeed, as mentioned before, if the production cost is low, then using responsive pricing to clear the market may amplify the double marginalization problem, and as a result the retailer could get hurt. When the production cost increases to $\tilde{C} < c < c_1^{DR}$, however, scenario OR becomes the best for the retailer: The retailer changes from payment postponement to pricing postponement. In this case, the production inflation gets more costly for the supplier and the increased wholesale price would dissuade the retailer from choosing a pay-by-delivery contract. Instead, the retailer will postpone pricing to transfer the risk to consumers. As a matter of fact, Proposition 2 implies that as long as $c > \tilde{C}$, the retailer will use responsive pricing. Hence, for the retailer, the benefit of matching supply with demand in the end market can always overcome the double marginalization problem when production cost is not low.

For production cost $c_1^{DR} < c < c_2^{DR}$, Proposition 2 shows a strict dominance relationship and the retailer achieves the highest profit in scenario DR, where both payment and retail pricing are postponed. Here, we observe a range of production cost over which using both postponement strategies dominates using just one of them. Indeed, as previously discussed, the benefit of using responsive pricing (i.e., more efficient retail market) is growing as c increases (i.e., less efficient production); and on the other hand, the retailer gets less benefit from a pay-by-delivery contract as c increases because production inflation becomes costly for the supplier. As a result, the net benefit achieves the highest at some medium range production cost. Note that when $c_2^{DR} < c < \frac{a}{b}\mu$, the supplier does not inflate production in scenario DR (i.e., Lemma 4), and thus we have $\pi_R^{OR} = \pi_R^{DR}$. Hence, if the production is so costly that inflation is not worthwhile, then the retailer might just prefer scenario OR instead.

Practical Implications: Since the postponement strategies are used to deal with the yield risk, one may intuit that using at least one of these strategies may help the retailer gain more profit; moreover, they are more likely to be adopted when the supplier is more unreliable. However, our findings bring cautions to such a conclusion. Specifically, Proposition 1 implies that even when the yield loss occurs with a relatively high likelihood, the retailer is still better off not using either postponement strategy if the upstream supplier's production is not costly. On the contrary, when full yield is very likely, the retailer should use some postponement strategy, as shown by Proposition 2(i). As such, the distribution of the random yield has an impact on the retailer's strategy selection in practice. These non-intuitive results are mainly due to the decentralized setting, where the unreliable yield causes the supplier to respond more aggressively to the retailer's postponement decisions, making them less effective for the retailer. Therefore, an unhealthy vertical relationship along the supply chain may render the yield risk management ineffective.

4.2. Strategic Interplay

We now explore the strategic interplay between the two postponement strategies. In particular, we examine whether these strategies are substitutes or complements. We define the two strategies to be substitutes [complements] if one strategy reduces [increases] the value of using the other.

Definition 1. The payment postponement strategy and the pricing postponement strategy are substitutes [complements] if $V < 0$ [$V > 0$], where $V = (\pi_R^{DR} - \pi_R^{OR}) - (\pi_R^{DA} - \pi_R^{OA})$.

The purpose of studying the strategic interplay is to infer when the two postponement strategies would make each other more effective. Positioned between the supplier and consumers in a supply chain, the retailer has several risk-mitigating tools at its disposal to shift the supply risk either upstream or downstream. Intuition suggests that these two directions of transferring risk are substitutes. However, can the two postponement strategies ever be complements? Will the effectiveness of one strategy be enhanced by using the other strategy? Since the indicator V is a function of production cost and the yield distribution, the answer to the above question depends on those parameters. In the following, we focus on a particular region of (α,θ) to determine the sign of V for all feasible production cost c.

Proposition 3. *Suppose $\alpha \geq 9/16$ and $(\alpha,\theta)\in\mathcal{H}_3$. Then, there exists a $0<\hat{C}<\frac{a}{b}\mu$ such that*
(i) $V<0$ for $0<c<c_1^{DR}$ or $\hat{C}<c<c^{DA}$;
(ii) $V>0$ for $c_1^{DR}<c<\hat{C}$;
(iii) $V=0$ for $c^{DA}<c<\frac{a}{b}\mu$.

First of all, note that when the cost is high, i.e., $c^{DA}<c<\frac{a}{b}\mu$, the supplier will never inflate production under a pay-by-delivery contract (see Lemmas 3 and 4). Hence, the payment postponement is actually irrelevant. As a result, it has no impact on the effectiveness of the pricing postponement strategy, i.e., $V=0$. Second, we confirm the intuition that the two strategies are indeed substitutes, but only when the production cost is either relatively low ($0<c<c_1^{DR}$) or relatively high ($\hat{C}<c<c^{DA}$). In these cases, we have shown in the previous results that either the responsive pricing or the pay-by-delivery contract may not be effective for the retailer with small or large c. Here, we further show that the effectiveness of one strategy could actually decrease when the other one is also in use.

More importantly, Proposition 3 also uncovers an interesting case in which the two postponement strategies are complements to each other: $V>0$ when $c_1^{DR}<c<\hat{C}$. That is, with medium range of production cost, both postponement strategies are beneficial to the retailer (see Proposition 2); moreover, the use of one strategy will enhance the effectiveness of the other strategy. Note that the above results hold for certain yield distribution. Our premise is that the supply is reliable in that full delivery is likely (i.e., large θ) and possible yield loss is small (i.e., large α). The more general situation will be numerically studied in Section 6.2.

<u>*Practical Implications:*</u> While the retailer's preference towards the use of various postponement strategies highlights their first-order differences, the sign of V studied here represents the difference in difference of postponing payment and postponing pricing. Hence, Proposition 3 complements

Propositions 1 and 2 by characterizing the second order interaction between the two strategies. Moreover, in practice, the interplay between the two postponement decisions does have some implications on the comparison between the sole impact and joint impact of the strategies. For example, suppose that the two strategies are substitutes, i.e., $V<0$, then $\pi_R^{DR}<\max\{\pi_R^{DA},\pi_R^{OR}\}$ if $\pi_R^{OA}>\min\{\pi_R^{DA},\pi_R^{OR}\}$. That is, if the status quo profit already dominates one of the strategies, then it is never good for the retailer to use both strategies. Conversely, suppose that the two strategies are complements, i.e., $V>0$, then $\pi_R^{DR}>\max\{\pi_R^{DA},\pi_R^{OR}\}$ if $\pi_R^{OA}<\min\{\pi_R^{DA},\pi_R^{OR}\}$. That is, if the retailer should take at least one of the postponement strategies to increase the current profit, then it is even better to use both strategies. These practical implications provide helpful guidance for the retailer to evaluate the benefit of the postponement strategies.

5. Channel Implications

In this section, we turn to investigate the systematic impact of a certain adopted postponement strategy on the retailer's profit, the upstream supplier's profit, and the downstream consumer surplus. Specifically, fixing the pricing scheme [procurement contract], we compare the profits and consumer surplus changes due to the use of payment postponement [pricing postponement]. Given that the retailer's profit gets improved, will the upstream supplier and the downstream consumers be positively or negatively affected? When will the entire supply chain benefit from either postponement strategy? These are the questions we attempt to answer.

5.1. The Impact of Payment Postponement

We begin with the study on the system-wide impact of the payment postponement strategy. When the retailer's pricing scheme is fixed (either ex ante or responsive pricing), the retailer choosing pay-by-delivery over pay-by-order will receive different reactions from the supplier, which will further affect the consumer surplus due to possible supply-demand mismatch. As a potential means of risk management, the pay-by-delivery strategy may help the retailer avoid directly dealing with the yield uncertainty by possible order inflation; rather, the upstream supplier may have to inflate production to cope with the risk now. The following proposition shows, for each pricing scheme, how this upward shift of risk-allocation influences each party along the supply chain.

PROPOSITION 4. ***(A)*** *Suppose that the retailer uses ex ante pricing. Then, there exists a* $0<\bar{C}<c^{DA}$ *such that: (i) Both the supplier's and the retailer's profits as well as the consumer surplus increase for* $\max\{0,c_2^{OA}\}<c<\max\{\bar{C},c_2^{OA}\}$*. (ii) The retailer's profit and the supplier's profit increase, but the consumer surplus decreases, for* $\max\{\bar{C},c_2^{OA}\}<c<c^{DA}$*. (iii) The supply chain is unaffected for* $0<c<\max\{0,c_1^{OA}\}$ *or* $c^{DA}<c<\frac{a}{b}\mu$*.*

(***R***) *Suppose that the retailer uses responsive pricing and* $\alpha \geq 9/16$*. Then we have: (i) Both the supplier's and the retailer's profits as well as the consumer surplus increase for* $0 < c < \tilde{C}$*. (ii) Both the supplier's and the retailer's profits as well as the consumer surplus decrease for* $\tilde{C} < c < c_1^{DR}$*. (iii) The retailer's profit and the consumer surplus increase, but the supplier's profit decreases, for* $c_1^{DR} < c < c_2^{DR}$*. (iv) The supply chain is unaffected for* $c_2^{DR} < c < \frac{a}{b}\mu$*.*

According to the proof of Lemma 1 (see Appendix B.1) the cost thresholds c_1^{OA} and c_2^{OA} may be negative, depending on the yield distribution. In particular, if $(\alpha,\theta) \in \mathcal{H}_3$, which includes relatively large θ, then $c_1^{OA} < c_2^{OA} < 0$. Thus, Proposition 4(**A**) becomes especially clean in this case: when the ex ante pricing scheme is adopted, the retailer choosing to postpone payment has the channel effect of (W,W,W) if the production cost $0 < c < \bar{C}$, and (W,W,L) if $\bar{C} < c < c^{DA}$. Here, the three items in the triplet, from left to right, denote the impact on the retailer's profit, the supplier's profit, and the consumer surplus. Specifically, "W" (win) denotes an increase in the profit or surplus due to the use of a postponement strategy and "L" (lose) denotes a decrease.

Therefore, the use of pay-by-delivery contract instead of pay-by-order can have a positive effect to all parties along the supply chain. Moreover, as shown above, when full yield occurs with relatively high probability (i.e., θ large), the (W,W,W) region is the largest ($0 < c < \bar{C}$). When the production cost is in the medium range (i.e., $\bar{C} < c < c^{DA}$), however, the payment postponement will only hurt the consumers. This result shows that, with relatively unlikely yield loss, letting the supplier bear its own yield risk is by large beneficial to the whole supply chain if the production cost is low. High production cost, on the other hand, will make the production inflation expensive, and is therefore passed by the firms to the end market, which ultimately hurts the consumers. If production is too costly for the supplier to consider inflation, i.e., $c^{DA} < c < \frac{a}{b}\mu$, then the payment postponement will not have any impact on the supply chain.

For general $(\alpha,\theta) \notin \mathcal{H}_3$, Proposition 4(**A**) also shows that the supply chain is unaffected when $0 < c < \max\{0, c_1^{OA}\}$. This situation occurs when the status quo scenario OA is already optimal (see the discussions following Proposition 1) for the retailer. Furthermore, it is worth noting from the proof in Appendix C that the supplier's profit never decreases when the scenario changes from OA to DA. Thus, interestingly, the retailer's payment postponement strategy, which shifts the yield risk back to the supplier, actually does not make the supplier worse off in this case.

Proposition 4(**R**) characterizes the channel implication of payment postponement when the retail price is decided in a responsive manner. Here, the overall positive impact (W,W,W) is observed when the production cost is not high, i.e., $0 < c < \tilde{C}$. As the production cost increases ($\tilde{C} < c < c_1^{DR}$), the retailer may be worse off from using a pay-by-delivery contract. Note that when this happens, everyone in the supply chain is hurt. That is, the postponement strategies may backfire and make the entire channel suffer. An interesting case arises when $c_1^{DR} < c < c_2^{DR}$, where we observe a (W,L,W) outcome. Recall that Lemma 4 shows that the supplier inflates the production level in this region. Hence, due to payment postponement, the supplier has to bear the yield risk but does not charge a higher wholesale price, which lowers its profit; but the consumers could benefit from the inflated production quantity and the responsive pricing.

Practical Implications: Proposition 4(**A**) and (**R**) together reveal the systematic impacts of the payment postponement strategy under the two pricing schemes. The foremost managerial implication is that, regardless of the pricing scheme, using a pay-by-delivery contract to shift the supply risk from the retailer to the upstream supplier could be Pareto optimal to the entire supply chain, especially when the production is not costly and the order is likely to be fully delivered. This provides an aligned incentive for every channel member to welcome a pay-by-delivery contract, which is usually shorter and easier to write (MacDonald et al. 2004). Secondly, our findings also have practical implications for the non-Pareto cases. Payment postponement strategy, when benefiting the retailer, may have a negative impact on either the supplier or the consumers, but not both, depending on which pricing scheme is in use. Specifically, consumers may get hurt under ex ante pricing whereas the supplier's profit may drop under responsive pricing. Therefore, although responsive pricing can protect the consumers by better clearing the end market, the retailer needs to consider using ex ante pricing to persuade the supplier to sign a pay-by-delivery contract.

5.2. The Impact of Pricing Postponement

Next, we turn to study the impact of pricing postponement on the supply chain. Compared to ex ante pricing, deciding retail price after yield realization allows the retailer to better match supply with demand in the end market. Thus, the yield risk exposed to the retailer is mitigated in the sense that consumers are sharing some of the risk as well. Such a downward transfer of the supply risk by the retailer generates systematic consequences in the equilibrium that affect all parties along the supply chain. Given the procurement contract (either pay-by-order or pay-by-delivery), the following proposition identifies and summarizes some of the interesting consequences caused by pricing postponement.

PROPOSITION 5. ***(O)*** *Suppose that a pay-by-order contract is in use. Then we have: (i) Both the supplier's and the retailer's profits as well as the consumer surplus increase for* $\max\{c^{OR}, c_2^{OA}, 0\} < c < \frac{a}{b}\mu$. *(ii) There exists an increasing function* $\Theta(\alpha)$ *such that, if* $0 < \theta < \Theta(\alpha)$, *then* $c_1^{OA} > \tilde{C} > 0$; *moreover, both the supplier's and the retailer's profits as well as the consumer surplus increase for*

$\tilde{C} < c < c_1^{OA}$. (iii) If $c^{OR} > 0$, then the retailer's profit and the consumer surplus increase, but the supplier's profit decreases, for $0 < c < \min\{c^{OR}, c_1^{OA}\}$.

***(D)** Suppose that a pay-by-delivery contract is in use and $\alpha \geq 9/16$. Then we have: (i) Both the supplier's and the retailer's profits as well as the consumer surplus increase for $c_1^{DR} < c < \frac{a}{b}\mu$. (ii) The supply chain is unaffected for $0 < c < c_1^{DR}$.*

Proposition 5(**O**) identifies conditions for two possible channel impacts of retail pricing postponement, namely, (W,W,W) and (W,L,W). Specifically, using responsive pricing rather than ex ante pricing could be a Pareto optimal move for the entire supply chain, regardless of the production cost. To see this, note that if the yield distribution satisfies $(\alpha, \theta) \in \mathcal{H}_3$, then both c^{OR} and c_2^{OA} are negative (see the proofs in Appendices B.1 and B.2). Hence, provided that the probability for yield loss is relatively small (i.e., large θ), Proposition 5(**O**)(i) states that for any $0 < c < \frac{a}{b}\mu$, pricing postponement simultaneously benefits the supplier, the retailer and the consumers alike.

When θ is small or medium, the yield is likely to suffer from proportional loss. However, responsive pricing can still achieve Pareto improvement for the supply chain. Indeed, a (W,W,W) outcome occurs for high or intermediate production cost. For example, we find situations where c needs to be larger than c^{OR} and c_2^{OA} when either or both of them are positive so that the entire channel may benefit from the pricing postponement; in addition, we also find that, when $0 < \theta < \Theta(\alpha)$, the same result holds for intermediate level of production cost, i.e., $\tilde{C} < c < c_1^{OA}$ (threshold $\tilde{C}$ was introduced in Proposition 1). Such results are consistent with our previous remark that the benefit of responsive pricing is increasing in c, when the double marginalization issue becomes light.

Another interesting result revealed by Proposition 5(**O**) is that the supplier could be hurt by the retailer's pricing postponement when the production cost is low. Such a (W,L,W) outcome is possible when $0 < c < \min\{c^{OR}, c_1^{OA}\}$ and the yield loss is both very severe (i.e., small α) and very likely (i.e., small θ). In this case, $c^{OR} > 0$ and the equilibrium result in scenario OR has two pieces; over the first piece, the retailer will inflate its order quantity (see Lemma 2). Hence, the retailer could order a large quantity and get away with a low wholesale price because the production is inexpensive. On the other hand, as the potential yield loss and the supply risk are too high under this circumstance, using responsive pricing can protect the retailer and the consumers to a large extent. Therefore, the supplier has to bear most of the negative impact of its own unreliability.

Proposition 5(**D**) further depicts the channel impact of responsive pricing under a pay-by-delivery contract. Under the assumption $\alpha \geq 9/16$, i.e., the yield loss is not severe, the result is very positive: The entire supply chain is either unaffected or strictly better off under responsive pricing. Moreover, we can see that the (W,W,W) region is for high production cost, $c_1^{DR} < c < \frac{a}{b}\mu$. Outside this range, the supplier chooses to inflate the production due to low cost, adding some surcharge to the retailer in the wholesale price. As a result of such a strategic interaction between the firms, the benefit of responsive pricing is partially offset, leaving all profits and surplus in the channel unchanged. However, with a high production cost, pricing postponement strictly benefits all parties when a pay-by-delivery contract is in use.

Practical Implications: The results in Proposition 5 clearly imply that the system-wide benefit of pricing postponement is reinforced by the use of payment postponement: To wit, if the retailer does not shift the yield risk to the upstream supplier, its responsive pricing strategy can hurt the supplier. Rather, given that the retailer does not have to bear the risk from the random yield, postponing the decision on the retail price will never hurt anyone along the supply chain. Moreover, under both payment schemes, the (W,W,W) situation appears when c is not small. Hence, our findings have another practical implication that the performance of a cost-inefficient decentralized supply chain is more likely to be improved by the retailer's strategic pricing postponement. Overall, the retailer's responsive pricing scheme may be relatively easy for all channel members to accept, especially when the supplier bears the yield risk and runs a costly production process.

6. Numerical Studies

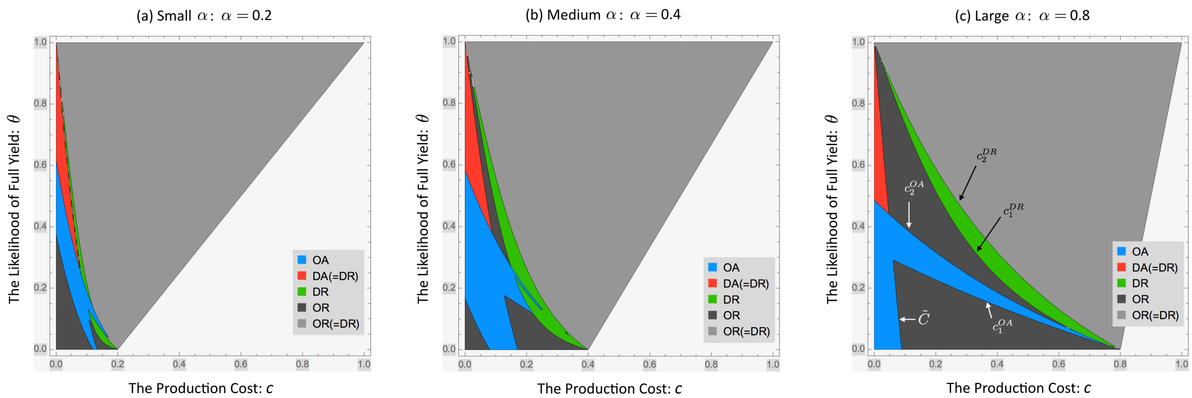
In the previous sections, we assume $\alpha \geq 9/16$ when considering scenario DR and examine specific regions such as $\mathcal{H}_i$ $(i=1,2,3)$ when scenario OA is involved. To confirm that these results continue to hold with general parameter choices, we now show the results of a series of extensive numerical studies, which not only echo our analytical results but also help us gain further insights. In each study, we fix $a=b=1$ and report the findings based on three representative α values, namely, $\alpha=0.2$ (small), $\alpha=0.4$ (medium), and $\alpha=0.8$ (large). The results are characterized as different regions on the feasible set $\{(c,\theta)\,|\,0<\theta<1, 0<c<\theta+\alpha(1-\theta)\}$.

6.1. Study 1: Retailer's Strategy Preference

To confirm the analytical results in Propositions 1 and 2, we use Figure 1 to depict the retailer's preferred scenario, $k^* = arg\max\{k|\pi_R^k\}$. The region labeled "OA" means the status quo scenario is already optimal for the retailer. According to Proposition 1, this occurs when the production cost is relatively small and $(\alpha,\theta)\in\mathcal{H}_1\cup\mathcal{H}_2$ (see the graphical illustration in Appendix B.1). Hence, Figure 1 here confirms that the above results largely remain valid with general parameters. Additionally, the numerical study also shows that the retailer prefers OA when $c_1^{OA}<c<c_2^{OA}$ (again, on $\mathcal{H}_1\cup\mathcal{H}_2$), which extends the analytical findings on the retailer's strategy selection.

Moreover, the retailer moves away from the status quo scenario when (c,θ) falls into other regions. Indeed, the scenarios that dominate others in different situations are given by Proposition

Figure 1 Strategy Preference: Which Strategy is Optimal for the Retailer?

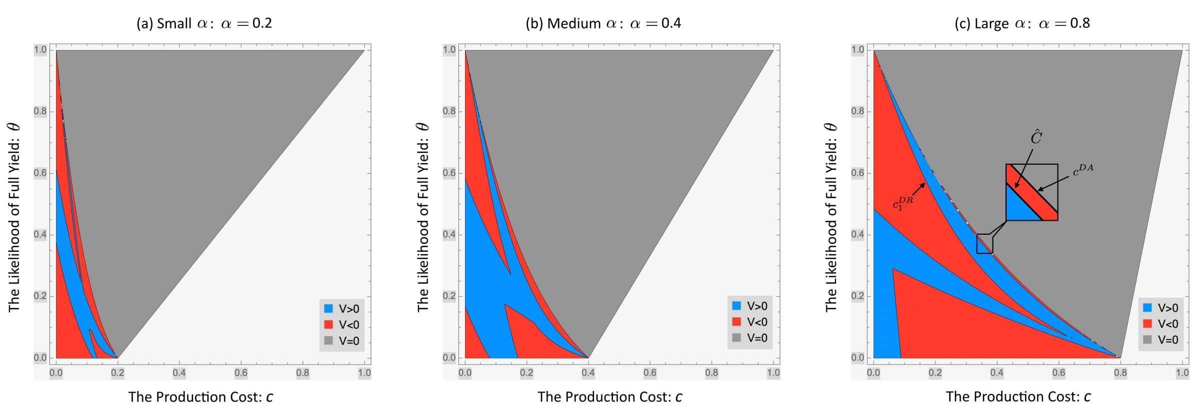


2. Although the proposition requires $\alpha \geq 9/16$, Figure 1 shows that the results still hold with small and medium α. Specifically, Figure 1 illustrates that payment postponement is preferred when c is small whereas pricing postponement is preferred when c is large; and for a certain medium range of production cost, using both postponement strategies is the best for the retailer.

6.2. Study 2: Interplay between the Two Postponement Strategies

The next numerical study focuses on the strategic interplay of the two postponement strategies, with the objective to confirm and generalize the results in Proposition 3. We look into the indicator V according to Definition 1 and characterize the respective regions of $V > 0$, $V < 0$, and $V = 0$ in Figure 2. While the analytical results given by Proposition 3 is based on special yield parameters

Figure 2 Strategy Interplay: Are They Substitutes or Complements?



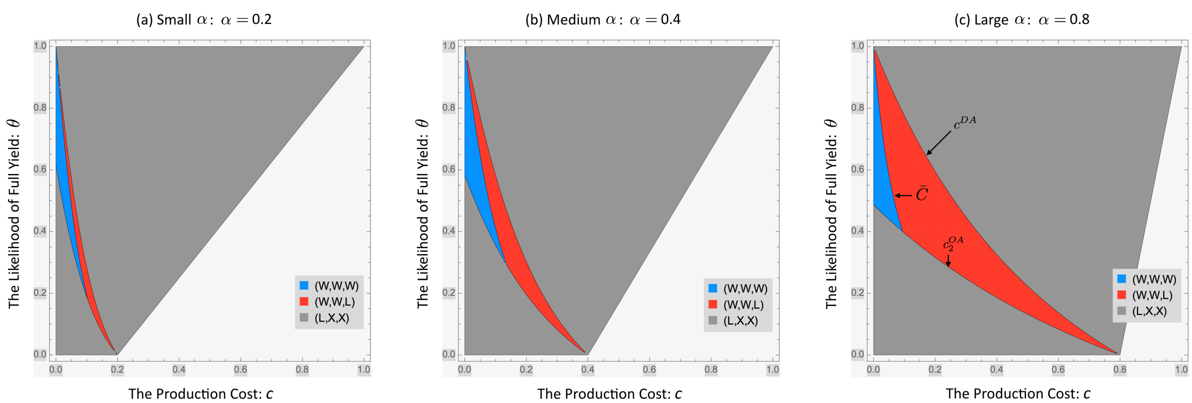
$\{\alpha \geq 9/16\} \cap \mathcal{H}_3$, the numerical study has supported the validity of our results in general settings. Specifically, when θ is large in Figure 2, we observe V changes its sign along the horizontal line in the pattern consistent with the insight of Proposition 3. In particular, for some medium range production cost c, payment postponement and pricing postponement are strategic complements.

Moreover, this range gets larger when α increases. Lastly, in addition to when θ is large, Figure 2 depicts another region where $V > 0$, i.e., when θ is medium or small. Comparing to Figure 1, we find that this region is exactly when OA dominates other scenarios, and therefore the retailer will never use either postponement strategy anyway.

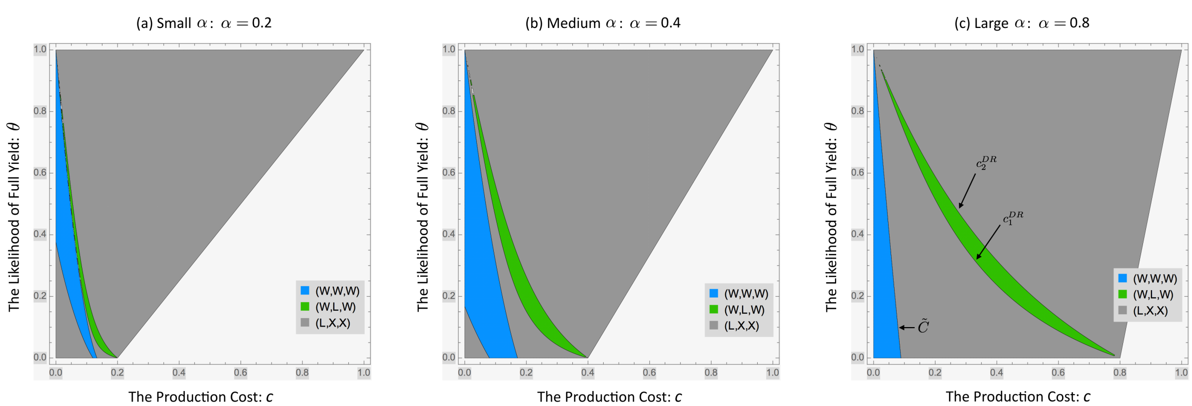
6.3. Study 3: Channel Impact of Payment Postponement

Next, we focus on the channel impact of payment postponement, which has been analytically examined by Proposition 4. First, Figure 3(**A**) confirms all our analytical results in Proposition 4(**A**): With relatively low production cost, payment postponement can benefit the entire supply chain; but with high production cost, the consumers may be worse off. Note that the region (L,X,X) means the retailer's profit is not *strictly* improved. We deem this situation as "lose" for the retailer and will not further consider the channel impact, which is denoted as "X".

Figure 3 **Channel Impact of Payment Postponement**



(a) (**A**)–Assume ex ante pricing and compare scenarios DA vs OA.



(b) (**R**)–Assume responsive pricing and compare scenarios DR vs OR.

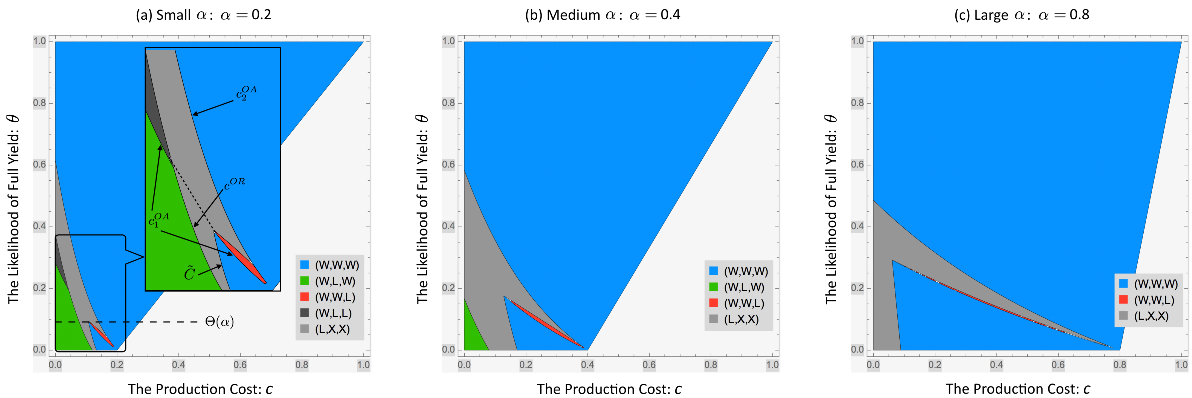
Second, under the responsive pricing scheme, Proposition 4(**R**) characterizes the channel impact when $\alpha \geq 9/16$. The analytical findings reveal that, when the retailer postpones payment, everyone in the supply chain benefits, except that the supplier is sometimes worse off. Moreover, (W,W,W) occurs when the production cost is low whereas (W,L,W) occurs when the cost is in the medium range. Figure 3(**R**)-(c) is a representative instance that demonstrates these results. Furthermore, Figure 3(**R**)-(a) and -(b) (small and medium α) uncover a situation that extends our analytical results. Here, the payment postponement strategy may not benefit the retailer when c and θ are small (see the bottom left corner in the graphs). In other words, when the yield is very likely to suffer from severe loss, the retailer who uses responsive pricing to mitigate the risk will not further use a pay-by-delivery contract to shift the risk to the upstream even if the production is inexpensive. Such a result can be explained by noting that the two strategies are substitutes in these regions; see Figure 2(a)&(b).

6.4. Study 4: Channel Impact of Pricing Postponement

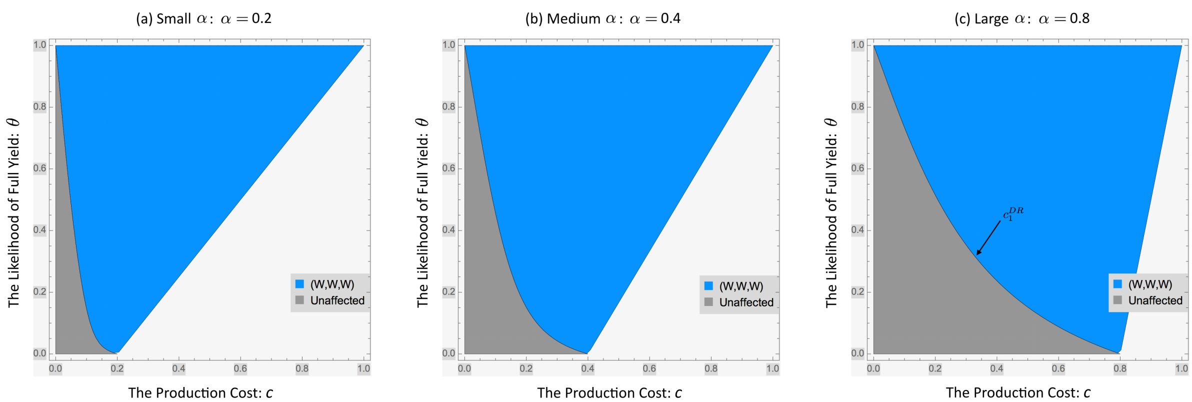
Lastly, we turn to the numerical study on the channel impact of pricing postponement. Figure 4 not only confirms the findings in Proposition 5, but also offers more insights. First, Figure 4(**O**) depicts the two regions (c is large, or c is medium but the yield loss is very likely) where pricing postponement is Pareto optimal, which have been analytically characterized in Proposition 5(**O**). Another confirmed analytical result is that pricing postponement may hurt the supplier, as we can observe the (W,L,W) region in Figure 4(**O**)-(a) and -(b). More interestingly, our numerical study uncovers a region of (W,L,L), which is not analytically captured. Specifically, when the retailer inflates the order to cope with the yield risk and uses responsive pricing to transfer the risk to downstream, both the supplier and the consumers may be worse off if the yield loss is both severe and likely; see Figure 4(**O**)-(a). Second, although the analytical result in Proposition 5(**D**) is obtained under the assumption $\alpha \geq 9/16$, the numerical results shown in Figure 4(**D**) indicate that it is valid for all α. The channel is either Pareto improved or completely unaffected by pricing postponement under a pay-by-delivery contract, and the smaller the potential yield fraction, the more likely that responsive pricing benefits the entire supply chain.

Practical Implications: We compare Figure 3 and Figure 4 to derive some additional insights. First, pricing postponement is more likely to enhance the overall channel performance than payment postponement, because the (W,W,W) region is much larger in Figure 4. Indeed, when the production is too costly, a pay-by-delivery contract is not able to incentivize the supplier to inflate the production and therefore has no effect on the channel. Second, despite the more frequent positive channel impact, pricing postponement may also result in the most undesirable channel consequence. In fact, the (W,L,L) case could occur only in Figure 4, rendering the retailer the only one to be better off. As such, our numerical studies reveal the contrasting features of the postponement strategies.

Figure 4 Channel Impact of Pricing Postponement



(a) (**O**)–Assume a pay-by-order contract and compare scenarios OR vs OA.



(b) (**D**)–Assume a pay-by-delivery contract and compare scenarios DR vs DA.

7. Extensions

7.1. Uniform Random Yield

The two-point distributed random yield in the main model is assumed so that useful insights could be derived with tractable analysis. To show the robustness of our findings, we look into uniform random yield as an extension to our model. Let $Y \sim U[x,1]$ be uniformly distributed over an interval $[x,1]$, where $x<1$. The support is chosen to capture the minimum yield guarantee, which is in parallel to α in the main model. An important difference that uniform random yield causes to our analysis is the supplier's optimal production level Q^* under the pay-by-delivery contract. Since the supplier's expected profit is $\mathbf{E}[w\hat{q} - cQ]$, where the delivered quantity $\hat{q} = \min\{YQ, q\}$, we maximize the profit function to obtain the optimal solution:

$$Q^* = \frac{q}{x_0}, \qquad \text{where } x_0 := \sqrt{x^2 + \frac{2c(1-x)}{w}} \in (x,1).$$

Recall that, in the main model where the yield fraction follows a two-point distribution, the supplier may not inflate the production level under the pay-by-delivery scheme if the cost is too high;

however, if it does, then the inflation will provide 100% safeguard against the fractional yield. By contrast, when the yield is uniformly distributed, the supplier *always* inflates the production when the retailer postpones the payment; however, a full delivery can never be guaranteed because $\hat{q} < q$ whenever $Y < x_0$. Therefore, we emphasize that the yield distribution is critical in our problem as it has a significant impact on our analysis as well as the structure of the equilibrium.

Recognizing the above difference in the supplier's production quantity with uniformly distributed yield, we apply the same analytical technique to solve for equilibrium outcomes for the four scenarios. Due to the continuous yield distribution assumption, the analysis becomes intractable when we solve the supplier's problem. Consequently, we resort to numerical studies and derive insights based on observations. In the following, we report our numerical findings and show that the managerial insights and the major takeaways from the main model are robust and generalizable when the yield fraction follows a uniform distribution. (The analytical details, the numerical setup, and the graphical exhibitions are given in Appendix D.)

(1) The retailer may already obtain the highest profit postponing neither pricing nor payment. Otherwise, the retailer tends to have higher profit when it uses responsive pricing (i.e., scenario DR and OR). Moreover, the two postponement strategies can be either substitutes or complements, depending on distribution parameters. In fact, we find the two strategies are more likely to be complements with the uniform yield than with the two-point yield. As such, our numerical study extends the findings in the main model regarding the interplay between the two strategies.

(2) Fixing the pricing strategy, the impact of payment postponement can be overall positive to the channel, although the (W,W,W) region is not large. Moreover, with the uniformly distributed yield, consumers may be worse off under ex ante pricing whereas the supplier may be worse off under responsive pricing. These results are consistent with what is shown in Figure 3. Fixing the payment strategy, postponing pricing is largely Pareto optimal for the supply chain, especially with the pay-by-delivery scheme. Under a pay-by-order contract, however, there are a large portion of cases where the consumers are worse off, which is a notable difference from Figure 4.

(3) Finally, we remark that the uniform distribution $U[x, 1]$ is in nature dissimilar to the two-point distribution. As a result, the numerical findings, although qualitatively resemble those for the main model, are still different in some aspects; see more details in Appendix D.

7.2. Incorporating Demand Risk

Given that the focus of our paper is on supply risk mitigation by postponement strategies, the main model assumes deterministic demand. However, in many practical settings, demand could be random as well. Therefore, in this extension, we bring our model closer to reality by incorporating demand risk. Recall that the deterministic demand in the main model takes the form of $d(p) = a - bp$. We now assume that the maximum market size $a = \hat{a}$ is a *random variable* independent of the yield fraction Y. Thus, the demand becomes $D(p) = \max\{\hat{a} - bp, 0\}$, which exhibits the risk of having a small market size realization. Note that only the retailer's optimization problem will be explicitly affected. Specifically, the retailer's revenue function $R(p, q|Y, \hat{a}) = p \min\{YQ, q, D(p)\}$, which is influenced by both supply risk and demand risk. To continue focusing on the supply risk, we assume that demand is realized *after* purchase and pricing decisions are made. In this way, the two postponement strategies, especially the responsive pricing scheme, are still aiming at mitigating the supply risk, and the results are comparable to those in the main model. Hence, given (w, Q), the retailer's optimization problem is

$$\begin{cases} \max_q \mathbf{E}_Y \left[\max_p \mathbf{E}_{\hat{a}} \left[R(p, q|Y, \hat{a})\right] - T^k(w, Q, q)\right] & \text{if} \quad k \in \{OR, DR\}; \\ \max_{q,p} \mathbf{E}_{Y,\hat{a}} \left[R(p, q|Y, \hat{a}) - T^k(w, Q, q)\right] & \text{if} \quad k \in \{OA, DA\}. \end{cases}$$

As indicated in the above formulation, even when the retailer adopts responsive pricing, only the yield risk is mitigated. Since the expectation with respect to the demand is taken before the price is optimized, the optimal responsive price $p^* = p^*(q, w, Q|Y)$, although contingent on Y, will still have to absorb the demand risk. Moreover, being convex, the demand function $D(p)$ can spoil the concavity of the retailer's revenue function, rendering significant analytical difficulties. We therefore use numerical approach to solve for the equilibrium in each scenario and compare the findings to the main model; the details of the analysis and the graphical exhibitions are given in Appendix E. Here, we present the relevant observations and discuss the robustness of our results.

(1) Consider the retailer's preference towards various scenarios and the interplay between the two postponement strategies. We observe that, the status quo scenario, OA, can already be optimal for the retailer. Moreover, there exists a considerable amount of instances where payment and pricing postponement are strategic complements, which echoes the key results from the main model. Therefore, these observations suggest that the main takeaways remain unchanged even when demand is random. However, as illustrated in Appendix E, the expectation taken with respect to $\hat{a}$ would have a considerable negative impact on the retailer's equilibrium profit.

(2) Consider the channel impacts of the postponement strategies. Overall, our observations for the results under demand risk are consistent with the main model; that is, either postponement strategy can be Pareto optimal. However, there are two major differences for the results in the presence of demand uncertainty. First, the (W,W,W) region becomes much smaller comparing to Figures 3&4. Second, in every pair of comparison, there are instances where either consumers or the supplier, or both, are hurt by the retailer's use of a certain postponement strategy.

(3) Overall, incorporating demand risk into our model does not qualitatively alter the main results from Section 6. However, we do observe that the random demand negatively affects the retailer's profit, and the use of the postponement strategies does not benefit the channel as much as when demand is deterministic. An important implication is that the demand risk, although directly facing the retailer, has its effect rippled through the supply chain, and the effectiveness of the (supply-risk-mitigating) postponement strategies could be weakened by the demand uncertainty.

7.3. Comparing with an Integrated Supply Chain

The wholesale price contract in the main model is widely used in practice and commonly assumed in research papers due to its simple and elegant form. However, this contract is known to induce the double marginalization issue and therefore fails to coordinate the supply chain. Since the vertical interaction between the supplier and the retailer is affected by the payment postponement strategy, would an altered timing structure of the price-only contract (i.e., the pay-by-delivery scheme) achieve channel coordination? To find out, let us consider a centralized setting, where the supplier is an internal production facility and the retailer is the sales department of the same firm. Under different pricing schemes, we solve for the total profit of such an integrated supply chain and compare it with the decentralized settings under different payment schemes.

Proposition 6. *Let Π^A and Π^R be the equilibrium profits of the integrated supply chain under ex ante and responsive pricing schemes, respectively. Then, for any $0 < \alpha, \theta < 1$ and production cost $0 < c < \frac{a}{b}\mu$, we have $\pi_R^k + \pi_S^k < \Pi^A$ for $k \in \{OA, DA\}$ and $\pi_R^k + \pi_S^k < \Pi^R$ for $k \in \{OR, DR\}$.*

Proposition 6 shows that, as long as the contract only specifies the unit wholesale price w, postponing the payment does not coordinate the supply chain. Moreover, this result is irrelevant of the timing of the retail price decision. We remark that, although price-only contract has been studied in various settings and its coordinating power has been examined under different assumptions, the timing structure of the contract is relatively under-explored. However, due to the uncertain supply, when and how the procurement payment is settled do affect the decisions of the channel members and thus should be considered as an important characteristic of the contract. Our findings fill in this gap to some extent. Together with Proposition 4, we show that, both being price-only wholesale contract, the pay-by-delivery scheme may indeed enhance the channel profit as compared to the pay-by-order scheme. This is the benefit of payment postponement, but this benefit mostly pertains to mitigating the supply risk. The double marginalization issue is still not addressed. As a result, stipulating the timing of the payment in a price-only contract does not eliminate the inefficiency, and the supply chain is not coordinated.

Lastly, Proposition 6 suggests that a two-part tariff agreement, where the wholesale price is set to be the production cost and the retailer makes a fixed transfer payment to the supplier, can achieve coordination. This is because (1) the supplier does not have any incentive to manipulate the production quantity, and (2) the retailer has incentive to order the same quantity as an integrated firm does to get the first-best profit. To find more sophisticated coordinating contracts in our setting, one has to cope with two difficulties. First, even though the quantity decision can be coordinated, the price decision is distorted under responsive pricing. In fact, in a newsvendor setting with price-dependent demand, Cachon (2003) shows that buy-back and quantity-flexibility contracts do not coordinate the channel, and revenue-sharing contract coordinates only under certain conditions. These contracts will run into the same trouble in our setting. Second, with random yield and the pay-by-delivery scheme, either supplier or retailer may inflate the quantity and their inflation incentive is particularly hard to coordinate (see, e.g., Tang and Kouvelis 2014). Therefore, in scenarios with payment postponement, the key is to coordinate the firms' quantity (inflation) decisions; e.g., the two-part tariff contract mentioned above takes care of this issue.

8. Conclusion

Firms are facing more risks in their global supply chains, and supply uncertainty is one of the most significant to combat against (Gyorey et al. 2010). Based on a parsimonious bilateral supply chain model, we study the risk mitigating strategies a retailer may adopt to mitigate the yield risk originated from the upstream production process. On the one hand, the retailer may specify the pay-by-delivery, instead of pay-by-order, payment scheme to delay the procurement payment until the delivered amount is known. In this way, the retailer makes the supplier bear some of the production yield risk. On the other hand, the retailer may decide the retail price after the yield is realized. Thus, the yield risk is partially transferred to the end market consumers.

Emphasizing the notion of postponement, the above strategies are both potentially effective in supply risk mitigation. However, the consequences of their interaction in a decentralized supply chain are not clear and worth investigating. For one thing, the combined effect of using both postponement strategies is not necessarily a simple addition of the individual effects. For another, the upstream supplier may take counteractive actions in response to the retailer's move, leaving the net result undetermined. Therefore, we study four scenarios depending on the retailer's strategy combinations, establish the unique equilibrium for the Stackelberg game in each scenario, and obtain the following findings and relevant managerial insights for practice.

First, it is important for the retailer to keep in mind that it may never strictly benefit from either postponement strategy, especially when the production cost is low and the yield is highly likely to suffer from loss. When using the postponement strategies can indeed improve the retailer's profit, the benefit of pricing postponement is increasing, whereas the benefit of payment postponement is decreasing, in the production cost. Second, contrary to intuition, we uncover a situation where postponing pricing and postponing payment are strategic complements for the retailer. Thus, the use of one strategy can increase the benefit of the other. Lastly, in terms of the channel implications, one important takeaway is that both postponement strategies may be beneficial to everyone along the supply chain. A pay-by-delivery contract has an overall positive impact on the channel when the production cost is low and the potential yield loss is not large. Responsive pricing may be Pareto optimal when the production cost is high and/or a pay-by-delivery contract is in use. In all, the above results could provide guidance to the retailer when it determines different payment and pricing schemes in a decentralized setting.

We conclude by pointing out a few interesting avenues for future research: (1) Other supply chain configurations are possible. For example, the retailer may maintain multiple potential suppliers in its supply base to take advantage of diversification. Analyzing different supply chain structures could lead to valuable insights. (2) If more powerful than the retailer, the supplier may be the one who decides payment postponement. Then, we have a grand game where each firm strategically chooses to postpone payment/pricing, which is worth studying in details. (3) In practice, if the marketing data does not support linear demand, then we need to assume other form of demand (e.g., MNL), which may give rise to different results and insights.

Acknowledgments

The authors thank the department editor, Professor Kamalini Ramdas, the anonymous associate editor, and two anonymous referees for their very helpful and constructive comments, which have led to significant improvements on both the content and the exposition of this study. G. Xiao acknowledges financial support from the Research Grants Council of Hong Kong [GRF Grant PolyU 15503920]. X. Guo acknowledges the support from the NationalNatural Science Foundation of China [Grant 72293564/72293560].

References

Anupindi, R., R. Akella. 1993. Diversification under supply uncertainty. *Management Science* **66**(8) 944–963.

Babich, V., C.S. Tang. 2012. Managing opportunistic supplier product adulteration: Deferred payments, inspection, and combined mechanisms. *Manufacturing & Service Operations Management* **14**(2) 301–314.

Cachon, Gérard P. 2003. Supply chain coordination with contracts. *Handbooks in operations research and management science* **11** 227–339.

Chen, J., D. Yao, S. Zheng. 2001. Optimal replenishment and rework with multiple unreliable supply sources. *Operations Research* **49**(3) 430–443.

Chick, E., H. Mamani, D. Simchi-Levi. 2008. Supply chain coordination and influenza vaccination. *Operations Research* **6**(56) 1493–1506.

Chod, J. 2017. Inventory, risk shifting, and trade credit. *Management Science* **63**(10) 3207–3225.

Chod, J., E. Lyandres, S.A. Yang. 2019. Trade credit and supplier competition. *Journal of Financial Economics* **131**(2) 484 – 505.

Chod, J., N. Rudi. 2005. Resource flexibility with responsive pricing. *Operations Research* **53**(3) 532–548.

Ciarallo, F., R. Akella, T. Morton. 1994. A periodic review, production planning model with uncertain capacity and uncertain demand—optimality of extended myopic policies. *Management Science* **40**(3) 320–332.

Cohen, M., G. Perakis, C. Thraves. 2018. A unifying framework for consumer surplus under demand uncertainty. *Working paper* .

Dada, M., N. Petruzzi, L. Schwarz. 2007. A newsvendor's procurement problem when suppliers are unreliable. *Mannufacturing & Service Operations Management* **1**(9) 9–32.

Dong, L., X. Geng, G. Xiao, N. Yang. 2022. Procurement strategies with unreliable suppliers under correlated random yields. *Manufacturing & Service Operations Management* **24**(1) 179–195.

Dong, L., G. Xiao, N. Yang. 2016. Supply diversification under price-dependent demand and random yield. *Working paper* .

Etienne, X.L., M.L. Mallory, S.H. Irwin. 2017. Estimating the cost of pre-harvest forward contracting corn and soybeans in illinois before and after 2007. *Agribusiness* **33**(3) 358–377.

Federgruen, A., N. Yang. 2008. Selecting a portfolio of suppliers under demand and supply risks. *Operations Research* **56**(4) 916–936.

Federgruen, A., N. Yang. 2009. Optimal supply diversification under general supply risks. *Operations Research* **57**(6) 1451–1468.

Federgruen, A., N. Yang. 2011. Procurement strategies with unreliable suppliers. *Operations Research* **59**(4) 1033–1039.

Federgruen, A., N. Yang. 2014. Infinite horizon strategies for replenishment systems with a general pool of suppliers. *Operations Research* **62**(1) 141–159.

Granot, D., S. Yin. 2008. Price and order postponement in a decentralized newsvendor model with multiplicative and price-dependent demand. *Operations Research* **56**(1) 121–139.

Grosfeld-Nir, A., Y. Gerchak. 2004. Multiple lot sizing in production to order with random yields: review of recent advances. *Annals of Operations Research* **1**(126) 43–69.

Gyorey, T., M. Jochim, S. Norton. 2010. The challenges ahead for supply chains: Mckinsey global survey results. *McKinsey Quarterly*. (November).

Henig, M., Y. Gerchak. 1990. The structure of periodic review policies in the presence of random yield. *Operations Research* **38**(4) 634–643.

Hu, B., D. Kostamis. 2015. Managing supply disruptions when sourcing from reliable and unreliable suppliers. *Production and Operations Management* **24**(5) 808–820.

Kouvelis, P. 2015. Monsanto company: new production introduction for seed corn & supply chain challenges. *Case Study, Olin Business School, Washington University in St. Louis*. (December).

Kouvelis, P., G. Xiao, N. Yang. 2018. On the properties of yield distributions in random yield problems: Conditions, class of distributions and relevant applications. *Production and Operations Management* **17**(7).

Kouvelis, P., G. Xiao, N. Yang. 2021. Role of risk aversion in price postponement under supply random yield. *Management Science* **67**(8) 4826–4844.

Kouvelis, P., W. Zhao. 2012. Financing the newsvendor: Supplier vs. bank, and the structure of optimal trade credit contracts. *Operations Research* **60**(3) 566–580.

Li, Q., S. Zheng. 2006. Joint inventory replenishment and pricing control for systems with uncertain yield and demand. *Operations Research* **54**(4) 696 – 705.

Li, T., S. Sethi, J. Zhang. 2017. Mitigating supply uncertainty: The interplay between diversification and pricing. *Production and Operations Management* **26**(3) 369–388.

MacDonald, J. M., J. Perry, M. C. Ahearn, D. Banker, W. Chambers, C. Dimitri, N. Key, K. E. Nelson, L. W. Southard. 2004. Contracts, markets, and prices: Organizing the production and use of agricultural commodities. *USDA-ERS Agricultural Economic Report Number 837* .

Rui, H., G. Lai. 2015. Sourcing with deferred payment and inspection under supplier product adulteration risk. *Production and Operations Management* **24**(6) 934–946.

Scott, N. 2003. *Agribusiness and commodity risk: strategies and management*. Incisive RWG Ltd, U.K.

Sinofsky, S. 2019. Pre-announcing as a strategy. *Learning by Shipping* URL `tinyurl.com/pre-ann`.

Tang, C., R. Yin. 2007. Responsive pricing under supply uncertainty. *European Journal of Operational Research* **182**(1) 239–255.

Tang, C.S., S.A. Yang, J. Wu. 2018. Sourcing from suppliers with financial constraints and performance risk. *Manufacturing & Service Operations Management* **20**(1) 70–84.

Tang, Y., H. Gurnani, D. Gupta. 2014. Managing disruptions in decentralized supply chains with endogenous supply reliability. *Production and Operations Management* **7**(23) 1198–1211.

Tang, Y., P. Kouvelis. 2011. Supplier diversification strategies in the present of yield uncertainty and buyer competition. *Manufacturing and Service Operations Management* **4**(13) 439–451.

Tang, Y., P. Kouvelis. 2014. Pay-back-revenue-sharing contract in coordinating supply chains with random yield. *Production and Operations Management* **12**(23) 1059–1478.

Van Mieghem, J., M. Dada. 1999. Price versus production postponement: Capacity and competition. *Management Sci.* **45**(12) 1631 – 1649.

Wang, Y., W. Gilland, B. Tomlin. 2009. Mitigating supply risk: dual-sourcing or process improvements? *Mannufacturing Service Operations Management* **3**(12) 489–510.

Wang, Y., Y. Xiao, N. Yang. 2014. Improving reliability of a shared supplier with competition and spillovers. *European Journal of Operational Research* **236**(2) 499 – 510.

Yang, S.A., J.R. Birge. 2018. Trade credit, risk sharing, and inventory financing portfolios. *Management Science* **64**(8) 3667–3689.

Yano, C., H. Lee. 1995. Lot sizing with random yields: a review. *Operations Research* **2**(43) 311–333.

Appendix A: Summary of Notations

Table 2 List of Notations and Scenarios.

Notations	Definition
Y	Random yield factor with a two-point support $\{\alpha, 1\}$.
θ	The probability that $Y=1$; $0<\theta<1$.
α	One of the realizations of Y; $0<\alpha<1$.
μ	The expected value of Y, i.e., $\mu=\alpha(1-\theta)+\theta$.
c	The unit production cost.
a, b	The intercept and the slope of the linear demand function.
d	The demand function for the end market; i.e., $d(p)=a-bp$.
w, Q	Supplier's wholesale price and the production level, respectively.
q, p	Retailer's order quantity and the retail price, respectively.
$\hat{q}$	The delivered amount after production, i.e., $\hat{q}=\min\{YQ, q\}$.
π_S^k, π_R^k	Supplier's and retailer's equilibrium profits, respectively, under scenario $k \in \{OA, DA, OR, DR\}$.
CS^k	Consumer surplus at equilibrium under scenario $k \in \{OA, DA, OR, DR\}$.
Abbreviations	**Scenarios**
OA	Pay-by-*Order* contract with *ex ante* pricing scheme
OR	Pay-by-*Order* contract with *responsive* pricing scheme
DA	Pay-by-*Delivery* contract with *ex ante* pricing scheme
DR	Pay-by-*Delivery* contract with *responsive* pricing scheme

Note: Abbreviations of the scenarios are used as superscripts throughout this paper.

Appendix B: Derivation of the Equilibrium Results

Since all of our statements in this paper are directly based on the equilibrium results in each scenario, we now first show their derivation. For the four scenarios, we derive the equilibrium outcomes, based on which all our results in the paper can be directly proved. To start, we give the basic assumptions for the parameters and provide preliminary results as building blocks. We assume the following to ensure positive and profitable production and non-negative demand:

$$0<c<\frac{a}{b}\mu, \quad 0<p, w<\frac{a}{b}, \quad w>\begin{cases} c & \text{when pay-by-order;} \\ \frac{c}{\mu} & \text{when pay-by-delivery.} \end{cases} \tag{2}$$

Next, we find the supplier's production level Q given the retailer's order quantity. If the pay-by-order contract is used, then the supplier never inflates the production, i.e., $Q=q$. However, under a pay-by-delivery contract, due to random yield, the supplier has an incentive to inflate the production. In this case, the supplier's profit function of the retailer's ordered amount q and the wholesale price w. First, note that the supplier's profit function is

$$\pi_S(Q)=-cQ+\theta w \min\{Q, q\}+(1-\theta) w \min\{\alpha Q, q\},$$

which is a piecewise linear function with the kinks at q and q/α. Based on our assumption, the maximum is achieved at one of the kinks, depending on the slope of the second linear piece. Therefore, the production level and the delivered amount, $(Q^*(q, w), \hat{q}^*(q, w))$, are given by

$$(Q^*, \hat{q}^*)=\begin{cases} (q, Yq) & \text{if } \frac{c}{\mu} \leq w<\frac{c}{\alpha(1-\theta)}; \\ (\frac{q}{\alpha}, q) & \text{if } \frac{c}{\alpha(1-\theta)} \leq w<\frac{a}{b}. \end{cases} \tag{3}$$

Finally, we calculate the optimal retail price p under responsive pricing scheme. The price is contingent to the delivered quantity $\hat{q}$. We maximize the retailer's total revenue $p\min\{d(p),\hat{q}\}$, and find

$$p^*(\hat{q}) = \frac{a-\hat{q}}{b}\mathbf{1}_{\left\{0<\hat{q}<\frac{a}{2}\right\}} + \frac{a}{2b}\mathbf{1}_{\left\{\hat{q}>\frac{a}{2}\right\}}. \tag{4}$$

In the sequel, we use backward induction to solve the Stackelberg game in each scenario. See Table 1 for the sequence of events. For the ease of exposition, the superscript $k\in\{OA,OR,DA,DR\}$ is suppressed as long as the context is clear (i.e., notations are defined only for the analysis of the specific scenario).

B.1. Scenario OA: Pay-By-Order and Ex Ante Pricing

Retailer's Problem. Given the wholesale price w, the retailer wants to optimize its profit $\pi_R(q,p) = -wq+\theta p\min\{q,a-bp\}+(1-\theta)p\min\{\alpha q,a-bp\}$. We apply sequential optimization and first optimize over the order quantity q. The objective function can be written as a three-piece linear function with kinks $a-bp$ and $(a-bp)/\alpha$. Therefore, the optimal order quantity is $q^*(p) = (a-bp)\mathbf{1}_{\left\{\frac{w}{\mu}<p<\frac{w}{\alpha(1-\theta)}\right\}} + \frac{a-bp}{\alpha}\mathbf{1}_{\left\{\frac{w}{\alpha(1-\theta)}<p<\frac{a}{b}\right\}}$. The retailer does not inflate the order in the first case whereas it does in the second case. Now, substitute the above back to the objective function, we have the retailer's profit as a function of the retail price: $\pi_R(p) = f_1(p)\mathbf{1}_{\left\{\frac{w}{\mu}<p<\frac{w}{\alpha(1-\theta)}\right\}} + f_2(p)\mathbf{1}_{\left\{\frac{w}{\alpha(1-\theta)}<p<\frac{a}{b}\right\}}$, where $f_1(p) = (\mu p-w)(a-bp)$ and $f_2(p) = (p-w/\alpha)(a-bp)$. The centers of these two quadratic functions are $p_1 = (a/b+w/\mu)/2$ and $p_2 = (a/b+w/\alpha)/2$, respectively. Let $f_i^* = f_i(p_i)$ be the unconstrained optimum ($i=1,2$). To optimize the profit, we discuss the wholesale price on the range of $0<w<a/b$; we can verify that $p^* = p_2\mathbf{1}_{\{0<w<\hat{w}\}} + p_1\mathbf{1}_{\left\{\hat{w}<w<a\alpha(1-\theta)/b\right\}}$, where $\hat{w}$ equalizes f_1^* and f_2^*, i.e., $\hat{w} := \frac{a\alpha(\mu-\theta\sqrt{\mu})}{b(\alpha+\theta)}$. The optimal order quantity can be derived accordingly.

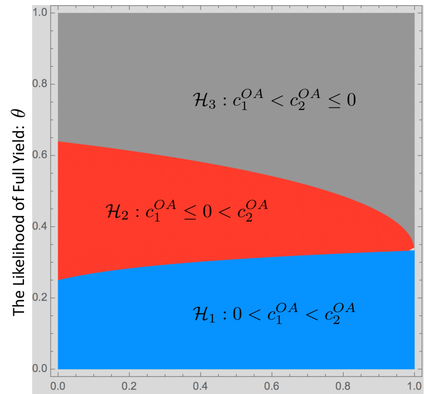
Supplier's Problem. Next, we solve for the supplier's optimal wholesale price. In scenario OA, its production level will be $Q=q$ and the delivered amount is $\hat{q}=Yq$. The objective function is then given by $\pi_S(w) = (w-c)q^*(w) = g_1(w)\mathbf{1}_{\{c<w<\hat{w}\}} + g_2(w)\mathbf{1}_{\left\{\hat{w}<w<\frac{a}{b}\mu\right\}}$, where $g_1(w) = (w-c)(a-bp_2(w))/\alpha$ and $g_2(w) = (w-c)(a-bp_1(w))$. Define their centers $w_1 = \frac{bc+a\alpha}{2b}$ and $w_2 = \frac{bc+a\mu}{2b}$; let $g_i^* = g_i(w_i)$ be the unconstrained optimum ($i=1,2$). Now, we discuss the production cost c over the range of $0<c<\frac{a}{b}\mu$ and we can find two critical thresholds c_1 and c_2, such that the optimal wholesale price (1) $w^*=w_1$ if $0<c<c_1$; (2) $w^*=\hat{w}$ if $c_1<c<c_2$; and (3) $w^*=w_2$ if $c_2<c<\frac{a}{b}\mu$. Here, let us formally define the cost thresholds

$$c_1^{OA} = c_1 := \frac{a\alpha(\mu-\alpha\theta-2\theta\sqrt{\mu})}{b(\alpha+\theta)}$$

and $c_2^{OA}=c_2$ is the second root of the polynomial $b^4\alpha^2(\alpha+\theta)^2x^4 - 4ab^3\alpha^2(\alpha^2-\theta^2)\mu x^3 + 2a^2b^2\mu^2\varphi_1(\alpha,\theta)x^2 - 4a^3b\alpha\mu^3(\alpha^3-4\alpha^2\theta-8\theta^2+11\alpha\theta^2)x + a^4\alpha^2\mu^3\varphi_2(\alpha,\theta)$, where $\varphi_1 = 3\alpha^4-6\alpha^3\theta+7\alpha^2\theta^2+8\alpha(-1+\theta)\theta^2-8\theta^3$ and $\varphi_2 = \alpha^3-5\alpha^2\theta-\alpha^3\theta-16\theta^2+19\alpha\theta^2+6\alpha^2\theta^2+25\theta^3-25\alpha\theta^3$. It is noteworthy that neither c_1^{OA} nor c_2^{OA} is guaranteed to be positive. In fact, there exist $0<\theta_1(\alpha)<\theta_2(\alpha)<1$ such that $0<c_1^{OA}<c_2^{OA}$ if $(\alpha,\theta)\in\mathcal{H}_1 := \{0\le\alpha\le1, 0<\theta<\theta_1\}$; $c_1^{OA}<0<c_2^{OA}$ if $(\alpha,\theta)\in\mathcal{H}_2 := \{0\le\alpha\le1, \theta_1<\theta<\theta_2\}$; and $c_1^{OA}<c_2^{OA}<0$ if $(\alpha,\theta)\in\mathcal{H}_3 := \{0\le\alpha\le1, \theta_2<\theta<1\}$. Here, $\theta_1(\alpha) = \frac{\left(1-4\alpha+\sqrt{1+8\alpha}\right)}{8(1-\alpha)}$ and $\theta_2(\alpha)$ is the third root of the polynomial $-\alpha^3+\alpha^2(5+\alpha)x+(16-19\alpha-6\alpha^2)x^2-25(1-\alpha)x^3$.

Lastly, Figure 5 below depicts the partition of the three sets mentioned in Lemma 1(i).

Figure 5 Illustration of the Partitioned Regions for Parameters (α, θ)



Summary. Let $[x]^+ = \min\{0, x\}$ and $\hat{w} = \frac{a\alpha(\mu - \theta\sqrt{\mu})}{b(\alpha+\theta)}$. Then, the firms' profits and the consumer surplus (obtained by using equations (1); same for other scenarios) at equilibrium, $(\pi_R^{OA}, \pi_S^{OA}, CS^{OA})$, are given as

$$\begin{cases} \left(\frac{(a\alpha-bc)^2}{16b\alpha^2}, \frac{(a\alpha-bc)^2}{8b\alpha^2}, \frac{(a\alpha-bc)^2}{32b\alpha^2}\right) & \text{if } \ 0 \leq c < [c_1^{OA}]^+ \\ \left(\frac{(a\alpha-b\hat{w})^2}{4b\alpha^2}, \frac{(a\alpha-b\hat{w})(\hat{w}-c)}{2\alpha^2}, \frac{(a\alpha-b\hat{w})^2}{8b\alpha^3}\right) & \text{if } \ [c_1^{OA}]^+ \leq c < [c_2^{OA}]^+ \\ \left(\frac{(a\mu-bc)^2}{16b\mu}, \frac{(a\mu-bc)^2}{8b\mu}, \frac{(a\mu-bc)^2(\alpha(2-\alpha)(1-\theta)+\theta)}{32b\mu^2}\right) & \text{if } \ [c_2^{OA}]^+ \leq c < \frac{a}{b}\mu, \end{cases}$$

B.2. Scenario OR: Pay-By-Order and Responsive Pricing

Retailer's Problem. Given the wholesale price w, we optimize the retailer's problem to find the optimal order quantity. The objective function is $\pi_R(q) = -wq + \mathbf{E}[p^*(\hat{q})\min\{\hat{q}, a - bp^*(\hat{q})\}]$. Moreover, the optimal retail price is always given by (4). Therefore, the profit function can be derived as $\pi_R(q) = h_1(q)\mathbf{1}_{\{0<q<\frac{a}{2}\}} + h_2(q)\mathbf{1}_{\{\frac{a}{2}<q<\frac{a}{2\alpha}\}} + \left(-wq + \frac{a^2}{4b}\right)\mathbf{1}_{\{\frac{a}{2\alpha}<q\}}$. The first two pieces are defined by $h_1(q) := -wq + \theta pq + (1-\theta)\frac{a-\alpha q}{b}\alpha q$ and $h_2(q) := -wq + \theta\frac{a^2}{4b} + (1-\theta)\frac{a-\alpha q}{b}\alpha q$. The centers are thus $q_1(w) = \frac{a\mu - bw}{2(\alpha^2(1-\theta)+\theta)}$ and $q_2(w) = \frac{a\alpha(1-\theta)-bw}{2\alpha^2(1-\theta)}$. Let $h_i^* = h_i(q_i)$ be the unconstrained optimum $(i = 1, 2)$.

We then discuss based on the wholesale price, and can find a critical threshold $\hat{w} = \frac{a}{b}\alpha(1-\alpha)(1-\theta) < \frac{a}{b}\mu$. We conclude that the sub-game optimal order quantity and profit functions of the retailer are $(q_2(w), h_2^*(w))$ if $0 < w < \hat{w}$; and $(q_1(w), h_1^*(w))$ if $\hat{w} < w < \frac{a}{b}\mu$

Supplier's Problem. Next, we move back to Stage 1 and solve the supplier's problem. Since its production level equals to the order quantity, the objective is simply to maximize the profit $\pi_S(w) = (w-c)q^*(w)$ by choosing a wholesale price from $c < w < \frac{a}{b}\mu$. Since the optimal order quantity has two pieces, we naturally write $\pi_S(w) = g_1(w)\mathbf{1}_{\{c<w<\hat{w}\}} + g_2(w)\mathbf{1}_{\{\hat{w}<w<\frac{a}{b}\mu\}}$, where the quadratic pieces $g_1(w) = (w-c)q_2(w)$ and $g_2(w) = (w-c)q_1(w)$, whose centers are denoted by w_1 and w_2, respectively. Let $g_i^* = g_i(w_i)$ be the unconstrained optimum $(i = 1, 2)$.

Then, we discuss c to determine the optimal solution for the supplier. We find that there exists a constant

$$c^{OR} := \frac{a}{b}\alpha(1-\theta)\left((1-\alpha) - \sqrt{\alpha^2 + \frac{\theta}{1-\theta}}\right) < \frac{a}{b}\mu,$$

such that $w^* = w_1$ if $0 < c < c^{OR}$ and $w^* = w_2$ otherwise. Note that c^{OR} may be negative, and we can verify that $c^{OR} \geq 0$ if and only if $0 < \alpha \leq \frac{1}{2}$ and $0 < \theta \leq \frac{1-2\alpha}{2(1-\alpha)}$. Hence, in a special case $\alpha > 1/2$, we have $c^{OR} < 0$ and thus there is only one expression for the equilibrium results.

Summary. The firms' profits and the consumer surplus at equilibrium, $(\pi_R^{OR}, \pi_S^{OR}, CS^{OR})$, are given as

$$\begin{cases} \left(\frac{(a\alpha(1-\theta)-bc)^2}{16b\alpha^2(1-\theta)} + \frac{a^2\theta}{4b}, \frac{(a\alpha(1-\theta)-bc)^2}{8b\alpha^2(1-\theta)}, \frac{(a\alpha(1-\theta)-bc)^2}{32b\alpha^2(1-\theta)} + \frac{a^2\theta}{8b}\right) & \text{if } \ 0 \leq c < [c^{OR}]^+ \\ \left(\frac{(a\mu-bc)^2}{16b(\alpha^2(1-\theta)+\theta)}, \frac{(a\mu-bc)^2}{8b(\alpha^2(1-\theta)+\theta)}, \frac{(a\mu-bc)^2}{32b(\alpha^2(1-\theta)+\theta)}\right) & \text{if } \ [c^{OR}]^+ \leq c < \frac{a}{b}\mu. \end{cases}$$

B.3. Scenario DA: Pay-By-Delivery and Ex Ante Pricing

Retailer's Problem. The retailer decides the order quantity and the retail price in this stage, which is again a two-step optimization problem. Under a pay-by-delivery contract, the retailer's profit function depends on the supplier's optimal production level given by (3). Hence, Let $w_0 := \frac{c}{\alpha(1-\theta)}$, then we have two cases: (1) if $w_0 < w < a/b$, then the delivered amount is $\hat{q} = q$ and $\pi_R(p,q) = -wq + p\min\{q, a-bp\}$, and the optimal outcomes are easily solved as $q^* = a - bp$ and $\pi_R(q^*, p) = (p-w)(a-bp) = f_1(p)/\mu$; (2) if $\frac{c}{\mu} < w < w_0$, then the delivered amount $\hat{q} = Yq$ and $\pi_R(p,q) = -\mu wq + \theta p \min\{q, a-bp\} + (1-\theta)p\min\{\alpha q, a-bp\}$, which has three linear pieces. So, in the second case, we define $p_0 := \frac{\mu w}{\alpha(1-\theta)}$ and then $q^*(p) = (a-bp)\mathbf{1}_{\{w<p<p_0\}} + \frac{a-bp}{\alpha}\mathbf{1}_{\{p_0<p<a/b\}}$ and $\pi_R(q^*,p) = f_1(p)\mathbf{1}_{\{w<p<p_0\}} + f_2(p)\mathbf{1}_{\{p_0<p<a/b\}}$. Here, $f_1(p) = \mu(p-w)(a-bp)$ and $f_2(p) = (p - \frac{\mu w}{\alpha})(a-bp)$ are two quadratic pieces. Let $p_1 = \frac{a+bw}{2b}$ and $p_2 = \frac{a+bw\mu/\alpha}{2b}$ be the centers and $f_i^* = f_i(p_i)$ be the unconstrained maximum, $i = 1, 2$.

In each case, we optimize the retailer's profit, and find that there exists a critical value for the wholesale price $\tilde{w} := \frac{a\alpha(1-\theta\sqrt{1/\mu})}{b(\alpha+\theta)}$ such that (1) if $c/\mu < w < \tilde{w}$, then $p^*(w) = p_2(w) = \frac{a+bw\mu/\alpha}{2b}$, $q^*(w) = \frac{a-bp_2(w)}{\alpha}$ (order inflated, production not inflated) and $\pi_R^*(w) = f_2^*(w)$; (2) if $\tilde{w} < w < w_0$, then $p^*(w) = p_1(w) = \frac{a+bw}{2b}$, $q^*(w) = a - bp_1(w)$ (no inflation) and $\pi_R^*(w) = f_1^*(w)$; (3) if $w_0 < w < a/b$, then $p^*(w) = p_1(w) = \frac{a+bw}{2b}$, $q^*(w) = a - bp_1(w)$ (order not inflated, production inflated) and $\pi_R^*(w) = f_1^*(w)/\mu$.

Supplier's Problem. The supplier's profit function is $\pi_S(w) = -cQ + \mathbf{E}[w\min\{YQ, q\}]$, which, based on the stage 2 results, can be rewritten as $\pi_S(w) = g_1(w)\mathbf{1}_{\left\{\frac{c}{\mu}<w<\tilde{w}\right\}} + g_2(w)\mathbf{1}_{\{\tilde{w}<w<w_0\}} + g_3(w)\mathbf{1}_{\left\{w_0<w<\frac{a}{b}\right\}}$. Here, $g_1(w) := \frac{b\mu^2}{2\alpha^2}(w - \frac{c}{\mu})(\frac{a\alpha}{b\mu} - w)$, $g_2(w) := \frac{b\mu}{2}(w - \frac{c}{\mu})(\frac{a}{b} - w)$ and $g_3(w) := \frac{b}{2}(w - \frac{c}{\alpha})(\frac{a}{b} - w)$; let w_i be the center of the quadratic piece $g_i(w)$, and $g_i^* = g_i(w_i)$ be the unconstrained maximum, $i = 1, 2, 3$. It is easy to verify the following useful properties regarding these pieces: (i) $w_1 < w_2 < w_3$; (ii) $g_1^* = g_3^*$; (iii) $g_1(\tilde{w}) > g_2(\tilde{w})$ and $g_2(w_0) = g_3(w_0)$.

Since the thresholds values of the wholesale price may not have a specific order, it is important we discuss the cases according to the order of the thresholds. Hence, by a discussion regarding the production cost, we find a critical value, namely,

$$c^{DA} = \frac{a\alpha(\mu - \theta\sqrt{\mu})}{b(\alpha+\theta)},$$

such that $w^* = w_3$ if $0 < c < c^{DA}$ and $w^* = w_2$ if $c^{DA} < c < \frac{a}{b}\mu$.

Summary. The firms' profits and the consumer surplus at equilibrium, $(\pi_R^{DA}, \pi_S^{DA}, CS^{DA})$, are given as

$$\begin{cases} \left(\frac{(bc-a\alpha)^2}{16b\alpha^2}, \frac{(bc-a\alpha)^2}{8b\alpha^2}, \frac{(bc-a\alpha)^2}{32b\alpha^2}\right) & \text{if } 0 \le c < c^{DA} \\ \left(\frac{(bc-a\mu)^2}{16b\mu}, \frac{(bc-a\mu)^2}{8b\mu}, \frac{(bc-a\mu)^2(\alpha(2-\alpha)(1-\theta)+\theta)}{32b\mu^2}\right) & \text{if } c^{DA} \le c < \frac{a}{b}\mu. \end{cases}$$

B.4. Scenario DR: Pay-By-Delivery and Responsive Pricing

Retailer's Problem. The retailer's profit function is $\pi_R(q) = \mathbf{E}\left[-w\hat{q} + p^*(\hat{q})\min\{\hat{q}, a - bp^*(\hat{q})\}\right]$, and it depends on the optimal pricing given by (4) and the supplier's delivery given by (3). Therefore, if $w_0 := \frac{c}{\alpha(1-\theta)} < w < \frac{a}{b}$, then $\hat{q} = q$ (supplier inflates production), and thus the profit function can be written as $-wq + \frac{a-q}{b}q$ if $0 < q < a/2$; and $-wq + \frac{a^2}{4b}$ if $a/2 < q$. Hence, when $w_0 < w < \frac{a}{b}$, the optimal order quantity is given by $q^*(w) = q_3(w) := \frac{a-bw}{2}$.

If $c/\mu < w < w_0$, then $\hat{q} = Yq$ (production not inflated), and thus $\pi_R(q)$ is a three-piece continuous function: $\pi_R(q) = h_1(q)\mathbf{1}_{\{0<q<\frac{a}{2}\}} + h_2(q)\mathbf{1}_{\{\frac{a}{2}<q<\frac{a}{2\alpha}\}} + (\frac{a^2}{4b} - \mu wq)\mathbf{1}_{\{\frac{a}{2\alpha}<q\}}$, where $h_1(q) := -\mu wq + \theta\frac{a-q}{b}q + (1-\theta)\frac{a-\alpha q}{b}\alpha q$ and $h_2(q) := -\mu wq + \theta\frac{a^2}{4b} + (1-\theta)\frac{a-\alpha q}{b}\alpha q$. Let q_i be the center of $h_i(q)$ and $h_i^* = h_i(q_i)$ be the unconstrained maximum, $i = 1, 2$; i.e., $q_1 = \frac{\mu(a-bw)}{2(\alpha^2(1-\theta)+\theta)}$ and $q_2 = \frac{a\alpha(1-\theta)-b\mu w}{2\alpha^2(1-\theta)}$. Solving the retailer's optimization problem, we obtain the solution $q^*(w) = q_2(w)\mathbf{1}_{\{\frac{c}{\mu}<w<\tilde{w}\}} + q_1(w)\mathbf{1}_{\{\tilde{w}<w<w_0\}} + q_3(w)\mathbf{1}_{\{w_0<w<\frac{a}{b}\}}$ where the critical wholesale price is given by $\tilde{w} := \frac{a\alpha(1-\alpha)(1-\theta)}{b\mu}$.

Supplier's Problem. Next, we maximize the supplier's profit function to find the optimal wholesale price. The objective function $\pi_S(w) = -cQ + \mathbf{E}[w\hat{q}]$ can be rewritten, based on the optimal order quantity from the last stage, as $\pi_S(w) = g_1(w)\mathbf{1}_{\{\frac{c}{\mu}<w<\tilde{w}\}} + g_2(w)\mathbf{1}_{\{\tilde{w}<w<w_0\}} + g_3(w)\mathbf{1}_{\{w_0<w<\frac{a}{b}\}}$, where $g_1(w) = (\mu w - c)q_2(w)$, $g_2(w) = (\mu w - c)q_1(w)$, and $g_3(w) = (w - \frac{c}{\alpha})q_3(w)$. Let w_i be the center of the quadratic piece $g_i(w)$ and $g_i^* = g_i(w_i)$ be the unconstrained maximum, $i = 1, 2, 3$. Notice that the profit function is not necessarily continuous at the kinks; so we need to consider boundary values when searching for the optimal solution.

Moreover, here we make an important assumption that the potential shortage of the yield is not too severe; specifically, $\alpha \ge 9/16$. Under this assumption, we may find two thresholds on the production costs, namely,

$$c_1^{DR} = \frac{a\alpha(1-\theta)\left(\mu^2 + \theta - 2\theta\sqrt{\alpha\mu(1-\alpha)(1-\theta)}\right)}{b\left(\alpha^2(1-\theta)^3 + 4\alpha\theta(1-\theta) + \theta(1+\theta)^2\right)}, \quad c_2^{DR} = \frac{a\alpha\mu(1-\theta)}{b(\mu+\theta)},$$

such that the optimal solution, (w^*, π_S^*), is simply given by (1) (w_3, g_3^*) if $0 < c < c_1^{DR}$; (2) $(w_0, g_2(w_0))$ if $c_1^{DR} < c < c_2^{DR}$; and (3) (w_2, g_2^*) if $c_2^{DR} < c < \frac{a}{b}\mu$.

Summary. The firms' profits and the consumer surplus at equilibrium, $(\pi_R^{DR}, \pi_S^{DR}, CS^{DR})$, are given as

$$\begin{cases} \left(\frac{(bc-a\alpha)^2}{16b\alpha^2}, \frac{(a\alpha-bc)^2}{8b\alpha^2}, \frac{(bc-a\alpha)^2}{32b\alpha^2}\right) & \text{if } 0 \le c < c_1^{DR} \\ \left(\frac{\mu^2(a\alpha(1-\theta)-bc)^2}{4b\alpha^2(1-\theta)^2(\alpha^2(1-\theta)+\theta)}, \frac{\mu\theta c(a\alpha(1-\theta)-bc)}{2\alpha^2(1-\theta)^2(\alpha^2(1-\theta)+\theta)}, \frac{\mu^2(a\alpha(1-\theta)-bc)^2}{8b\alpha^2(1-\theta)^2(\alpha^2(1-\theta)+\theta)^2}\right) & \text{if } c_1^{DR} \le c < c_2^{DR} \\ \left(\frac{(a\mu-bc)^2}{16b(\alpha^2(1-\theta)+\theta)}, \frac{(a\mu-bc)^2}{8b(\alpha^2(1-\theta)+\theta)}, \frac{(a\mu-bc)^2}{32b(\alpha^2(1-\theta)+\theta)}\right) & \text{if } c_2^{DR} \le c < \frac{a}{b}\mu. \end{cases}$$

B.5. Equilibrium Results for an Integrated Supply Chain

For a centralized supply chain, the decision variables are the production level, Q, and the retail price, p. Depending on the timing of the decision making, the price may be determined either prior to the yield realization (ex ante pricing) or after it (responsive pricing).

Ex Ante Pricing In this scenario, firm decides production level Q and retail price p before yield realizes, and the available-for-sale amount is $\hat{q} = YQ$. The firm's expected profit is therefore given by $\Pi(Q,p) = -cQ + \mathbf{E}[p\min\{YQ, a-bp\}]$, which is a piecewise linear function with respect to Q. Let $p_0 := \frac{c}{\alpha(1-\theta)}$. Then $Q^*(p) = a - bp$ if $c/\mu < p < p_0$ (No Inflation); and $Q^*(p) = \frac{a-bp}{\alpha}$ if $p_0 < p < a/b$ (Inflation). Substituting back to the profit function, we have $\Pi(p) = f_1(p)\mathbf{1}_{\{\frac{c}{\mu}<p<p_0\}} + f_2(p)\mathbf{1}_{\{p_0<p<\frac{a}{b}\}}$, where $f_1(p) := (a-bp)(\mu p - c)$ and $f_2(p) := (a-bp)(p - c/\alpha)$. It is straightforward to optimize the profit and obtain the critical production cost $c_A := \frac{a\alpha(\mu-\theta\sqrt{\mu})}{b(\alpha+\theta)}$; and the optimal profit is

$$\Pi^A = \frac{(a\alpha - bc)^2}{4b\alpha^2}\mathbf{1}_{\{0<c<c_A\}} + \frac{(a\mu - bc)^2}{4b\mu}\mathbf{1}_{\{c_A<c<\frac{a}{b}\mu\}}.$$

Responsive Pricing In this scenario, firm decides production level Q before yield realizes and it then sets the responsive price p, which is specified by (4). The firm's total profit can be written as $\Pi(Q) = -cQ + \mathbf{E}[p^*(YQ)\min\{YQ, a-bp^*(YQ)\}] = h_1(Q)\mathbf{1}_{\{0\le Q<\frac{a}{2}\}} + h_2(Q)\mathbf{1}_{\{\frac{a}{2}\le Q<\frac{a}{2\alpha}\}} + (-cQ + \frac{a^2}{4b})\mathbf{1}_{\{Q\ge\frac{1}{2\alpha}\}}$. Here, $h_1(Q) := Q\left(a\mu - bc - Q(\theta + \alpha^2(1-\theta))\right)/b$ and $h_2(Q) := \left(-\alpha^2(1-\theta)Q^2 + (a\alpha(1-\theta) - bc)Q + \theta\alpha^2/4\right)/b$. Since the profit function can be easily verified to be continuous, and its last piece is decreasing, we can focus on the first two pieces and identify the critical production cost $c_R := \frac{a\alpha(1-\alpha)(1-\theta))}{b}$ such that the optimal profit is

$$\Pi^R = \frac{(a\alpha - bc)^2}{4b\alpha^2}\mathbf{1}_{\{0<c<c_R\}} + \left(\frac{(a\alpha(1-\theta) - bc)^2}{4b\alpha^2(1-\theta)} + \frac{\theta a^2}{4b}\right)\mathbf{1}_{\{c_R<c<a\mu/b\}}.$$

Appendix C: Proofs of Statements

Based on the equilibrium results derived above, we can directly prove all our statements. Since the proofs are mostly straightforward comparisons, we just provide the key steps for the sake of exposition brevity. To further enhance the readability, we first list the three critical values of production cost that appear in the statements. First, $\tilde{C}$, which appears in Propositions 1, 2, 4(**R**) and 5(**O**), is given by

$$\tilde{C} := \frac{a\alpha(1 - \sqrt{\alpha^2(1-\theta) + \theta})}{b(1+\alpha)}. \tag{5}$$

Second, $\hat{C}$, which appears in Proposition 3, is given by

$$\hat{C} := \frac{a\alpha(1-\theta)}{b}\left(\frac{\phi_a - \theta\sqrt{\phi_c^{(1)} + \phi_c^{(2)} + \phi_c^{(3)}}}{\phi_b}\right). \tag{6}$$

The above parameters are further given by the following. Let $y := \alpha(1-\theta)$; then, $\phi_a = \alpha y^2(y + 2 - \theta) + 10\theta y - \theta(1-13\theta)y - \theta^2(1-5\theta)$, $\phi_b = y^3\alpha + y^3(2+\theta) + y^2\theta(11+\theta) - y\theta(1-14\theta+\theta^2) - \theta^2(1-6\theta+\theta^2)$, $\phi_c^{(1)} = y^4(1+y)\alpha^2$, $\phi_c^{(2)} = y^2(\theta(14-15\theta+5\theta^2) + (3-10\theta+20\theta^2)y - (1-11\theta)y^2)\alpha$, $\phi_c^{(3)} = 4\theta^5 - \theta^2(1-6\theta-15\theta^2)y - \theta(1-17\theta-16\theta^2)y^2$.

Third, $\bar{C}$, which appears in Proposition 4(**A**), is given by

$$\bar{C} := \frac{a\alpha}{b}\left(\frac{\varphi_a - \theta\mu\sqrt{\varphi_c}}{\varphi_b}\right), \tag{7}$$

where $\varphi_a = \alpha(1-\theta)(\alpha - \alpha^2(1-\theta) + 2\theta(1-\alpha)) + \theta^2$, $\varphi_b = \varphi_a + \alpha\theta\left(\alpha + \mu(1-\alpha)\right)$, $\phi_c = \alpha(1-\theta)(2-\alpha) + \theta$.

Proof of Lemmas 1-4. These statements follow directly from the results in Appendix B. Q.E.D.

Proof of Propositions 1 and 2. In these two propositions, we focus on the comparison of the retailer's profit. Since $\alpha \geq 9/16$, the threshold $c^{OR} < 0$ and thus only the second piece of π_R^{OR} is relevant. In addition, by direct comparison, we can order the thresholds as $c_1^{OA} < c_2^{OA} < c_2^{DR}$ and $c_1^{DR} < c_2^{DR} < c^{DA}$. The ordering is essential to our proof because we can identify the correct pieces of the retailer's profit function to compare.

Hence, on $\mathcal{H}_1$, $c_1^{OA} > 0$ and we check that $\pi_R^{OA} = \pi_R^{DA} = \pi_R^{DR}$ on the first piece. Moreover, $\pi_R^{OA} > \pi_R^{OR}$ on the first piece when $c < \tilde{C}$, where $\tilde{C}$ is given in (5). On $\mathcal{H}_2$, $c_1^{OA} < 0 < c_2^{OA}$ and we use the second piece of π_R^{OA} to compare; and then we find $\pi_R^{OA} > \pi_R^{DA} = \pi_R^{DR} > \pi_R^{OR}$ if $0 < c < \tilde{C}$. Furthermore, it is easy to compare the third piece of π_R^{OA} with the second piece of π_R^{OR}; so we find that $\pi_R^{OA} < \pi_R^{OR}$ if $\max\{0, c_2^{OA}\} < c < \frac{a}{b}\mu$. Finally, we may use the correct pieces of π_R^{OR}, π_R^{DA} and π_R^{DR} to straightforwardly compare them on $c_1^{DR} < c < c_2^{DR}$ and $c_2^{DR} < c < \frac{a}{b}\mu$, respectively. Q.E.D.

Proof of Proposition 3. By focusing on $\{\alpha \geq 9/16\} \cup \mathcal{H}_3$, the profit functions π_R^{OA} and π_R^{OR} both have only one piece relevant. Additionally, we have $c_1^{DR} < c_2^{DR} < c^{DA}$. Therefore, $V = (\pi_R^{DR} - \pi_R^{OR}) - (\pi_R^{DA} - \pi_R^{OA})$ has four pieces on different intervals; so we can readily check $V = 0$ for $c^{DA} < c < \frac{a}{b}\mu$ and $V < 0$ for $0 < c < c_1^{DR}$ and $c_2^{DR} < c < c^{DA}$. On the piece $c_1^{DR} < c < c_2^{DR}$, however, the sign of V depends on the production cost. Thus, we identify a critical value $\hat{C}$, which is given in (6), such that $c_1^{DR} < \hat{C} < c_2^{DR}$; moreover, $V > 0$ if $c_1^{DR} < c < \hat{C}$ and $V < 0$ if $\hat{C} < c \leq c_2^{DR}$. Q.E.D.

Proof of Proposition 4. Part (**A**) compares the firms' profits and consumer surplus between scenarios DA and OA. Note that $c_1^{OA} < c_2^{OA} < c^{DA}$, and thus there are four intervals, on which we can use the corresponding function piece to prove the statements. First, it is easy to verify that $\pi_R^{OA} = \pi_R^{DA}$, $\pi_S^{OA} = \pi_S^{DA}$ and $CS^{OA} = CS^{DA}$ for small or large production cost; i.e., $0 < c < \max\{0, c_1^{OA}\}$ or $c^{DA} < c < \frac{a}{b}\mu$. Second, it can be again directly verified that $\pi_S^{OA} < \pi_S^{DA}$ for all $\max\{0, c_1^{OA}\} < c < c^{DA}$, whereas $\pi_R^{OA} < \pi_R^{DA}$ only for $\max\{0, c_2^{OA}\} < c < c^{DA}$. Third, as for the consumer surplus, we find a critical value $\bar{C}$, which is given in (7), such that $CS^{OA} < CS^{DA}$ if $0 < c < \bar{C}$ and $c > \max\{0, c_2^{OA}\}$ (the third piece of equilibrium functions in scenario OA). Combining the above findings proves the statements.

Part (**R**) focuses on the comparison between scenarios DR and OR. Since $\alpha \geq 9/16$, the threshold $c^{OR} < 0$ and thus we look into the three intervals from scenario DR. Hence, statements (iii) and (iv) in the proposition are easily verified. As for the interval $0 < c < c_1^{DR}$, we find that the critical value $\tilde{C} \in (0, c_1^{DR})$, which is given in (5), serves as the breaking point that determines which scenario dominates the other. In particular, for $c < [>] \tilde{C}$, $\pi_R^{DR} > [<] \pi_R^{OR}$, $\pi_S^{DR} > [<] \pi_S^{OR}$ and $CS^{DR} > [<] CS^{OR}$. Q.E.D.

Proof of Proposition 5. Part (**O**) compares the equilibrium outcomes between scenarios OR and OA. To prove statement (i), note $\max\{c^{OR}, c_2^{OA}, 0\} < c < \frac{a}{b}\mu$, and thus we only need to compare the second piece of equilibrium functions in scenario OR and the third piece of those in scenarios OA. Similarly, to prove statement (iii), we compare the first piece of equilibrium functions in both scenarios because $0 < c < \min\{c^{OR}, c_1^{OA}\}$. The results are therefore immediate. Lastly, to verify statement (ii), we first introduce the threshold value $\Theta(\alpha)$.

$$\Theta(\alpha) := \frac{\alpha\left(3(1-\alpha-4\alpha^2) + 4\sqrt{\alpha(3+5\alpha+\alpha^2)}\right)}{(1-\alpha)(3+4\alpha)^2}.$$

Then, it follows that $c_1^{OA} > \tilde{C} > 0$ if $0 < \theta < \Theta(\alpha)$, where $\tilde{C}$ is given in (5). Furthermore, we have $c^{OR} < \tilde{C}$. Hence, for $\tilde{C} < c < c_1^{OA}$, we compare the second piece of equilibrium functions in scenario OR and the first piece of those in scenarios OA, and find that $\pi_R^{OA} < \pi_R^{OR}$, $\pi_S^{OA} < \pi_S^{OR}$ and $CS^{OA} < CS^{OR}$.

Part (**D**) assumes $\alpha \geq 9/16$ and focuses on scenarios DR and DA. Since $c_1^{DR} < c_2^{DR} < c^{DA}$, the comparison is straightforwardly divided into four intervals. Then, it is easily verified that the equilibrium outcomes stay the same if $0 < c < c_1^{DR}$, and the (W,W,W) situation occurs otherwise. Q.E.D.

Appendix D: Detailed Analysis for Uniform Random Yield

We show the analytical details and our numerical studies for Section 7.1, where the yield fraction is assumed to follow a uniform distribution over $[x, 1]$. Before the analysis, some observations are noteworthy.

First, in scenarios with responsive pricing, the optimal retail price is still given by (4); moreover, the revenue of the retailer is also contingent on the delivery and is given by

$$p^* \min\{a - bp^*, \hat{q}\} = \frac{(a-\hat{q})\hat{q}}{b}\mathbf{1}_{\{0<\hat{q}<\frac{a}{2}\}} + \frac{a^2}{4b}\mathbf{1}_{\{\hat{q}\geq\frac{a}{2}\}}. \tag{8}$$

Second, we can write out the consumer surplus given the equilibrium outcomes. Specifically, for ex ante pricing scenarios,

$$CS = \mathbf{E}\left[\left(\frac{a}{b} - p\right)Z - \frac{1}{2b}Z^2\right],$$

where $Z = \min\{a - bp, \hat{q}\}$ is the sales quantity. For responsive pricing scenarios, since we have the price as a function of delivery, the consumer surplus is

$$CS = \mathbf{E}\left[\frac{\hat{q}^2}{2b}\mathbf{1}_{\{0<\hat{q}<\frac{a}{2}\}} + \frac{a^2}{8b}\mathbf{1}_{\{\hat{q}\geq\frac{a}{2}\}}\right].$$

Lastly, we define $\hat{Y} = \min\{Y, x_0\}$; it is easy to see $\mathbf{E}\hat{Y} = \mu - \frac{(1-x_0)^2}{2(1-x)}$, where $\mu := (x+1)/2$.

In the following, we analyze the retailer's problem in each of the four scenarios. For the supplier's problem, since we have to use numerical method, we also briefly describe the procedures. Finally, the consumer surplus can be computed once we obtain the equilibrium order quantity and retail price, based on the formula we presented in the above. Like before, we omit the superscript when it is clear in the context.

Scenario OA: In this scenario, the supplier does not need to inflate production, and therefore $Q = q$ and the delivery $\hat{q} = Yq$. Given wholesale price w, the retailer's profit function is

$$\pi_R(p, q) = \mathbf{E}\left[p\min\{a - bp, Yq\} - wq\right].$$

We optimize the order quantity q over the interval $(a - bp, (a - bp)/x)$ (because it is easy to verify that the profit function is monotone outside the interval. We can check that the function is concave on the interval, with the first order condition (FOC) given by

$$\int_x^{\frac{a-bp}{q}} y dF(y) = \frac{w}{p}.$$

Define $\tilde{x} := \sqrt{x^2 + 2w(1-x)/p}$; then, the optimal order quantity as a function of price is $q^* = (a - bp)/\tilde{x}$. Next, optimize the retailer's profit (as a function of price p):

$$\pi_R(p) = \frac{a - bp}{\tilde{x}}\left(p\left(\mu - \frac{(1-\tilde{x})^2}{2(1-x)}\right) - w\right).$$

According to Kouvelis et al. (2018), the unique optimal price can be solved from the FOC efficiently. Hence, we apply bisect search over the interval $p \in [w/\mu, a/b]$ to find the optimal price p^*.

Now, given a, b, c, we can find the equilibrium outcome in the following procedure:

(1) For every $w \in (c, a\mu/b)$, find optimal price $p^*(w)$ from FOC based on the above discussion; then we also have $q^*(w)$.

(2) Search $\pi_S = (w-c)q^*(w)$ to find the optimal wholesale price w^*.

Scenario OR: Recall that the optimal price and revenue of the retailer is given by (4) and (8), respectively. In addition, the supplier does not inflate production level in this scenario; i.e., $Q = q$. Then, the retailer optimizes its expected profit as a function of order quantity:

$$\pi_R(q) = \mathbf{E}\left[\frac{(a-Yq)Yq}{b}\mathbf{1}_{\{Yq<a/2\}} + \frac{a^2}{4b}\mathbf{1}_{\{Yq\geq a/2\}}\right] - wq.$$

Consider two pieces: First, if $0 < q < a/2$, then $\pi_R(q) = \left(\frac{a}{b}\mu - w\right)q - \frac{\mathbf{E}Y^2}{b}q^2$ is a quadratic function; second, if $a/2 \leq q < a/(2x)$, then $\pi_R(q)$ is concave and the FOC is

$$\int_x^{\frac{a}{2q}} \frac{(a-yq)yq}{b} dF(y) = w.$$

Therefore, given a w, the optimal order quantity of the retailer, $q^*(w)$, can be computed efficiently (by considering the two pieces and comparing their optima). Note that we do not need to consider the case of larger q because the profit function is decreasing on that region.

Now, given a, b, c, we can find the equilibrium outcome in the following procedure:

(1) For every $w \in (c, a/b)$, find optimal price $q^*(w)$ of the retailer based on the above discussion.

(2) Search $\pi_S = (w-c)q^*(w)$ to find the optimal wholesale price w^*.

Scenario DA: Recall that the supplier will inflate the production level to be $Q = q/x_0$ given the order quantity. Similar to the previous scenario, the retailer maximizes its profit function over (p, q):

$$\pi_R(p,q) = \mathbf{E}\left[p\min\{a-bp, \hat{q}\} - w\hat{q}\right], \text{ where } \hat{q} = \frac{q}{x_0}\hat{Y}.$$

Recall $\hat{Y} = \min\{Y, x_0\}$. We first optimize order quantity, focusing on the interval $(a-bp, (a-bp)x_0/x)$ because the function is monotone otherwise. It is easy to verify that the function is concave and the FOC is

$$\int_x^{\frac{a-bp}{q/x_0}} y dF(y) = \frac{\hat{w}}{p}, \text{ where } \hat{w} = w\left(\mu - \frac{(1-x_0)^2}{2(1-x)}\right).$$

Define $\hat{x} := \sqrt{x^2 + 2\hat{w}(1-x)/p}$; then, the solution to the above FOC is $q = (a-bp)x_0/\hat{x}$. Note that this quantity is smaller than $(a-bp)x_0/x$, but not necessarily larger than $a-bp$. Hence, the optimal order quantity as a function of price should be

$$q^* = \frac{(a-bp)x_0}{\bar{x}}, \text{ where } \bar{x} = \min\{x_0, \hat{x}\}.$$

Next, we optimize the retailer's profit (as a function of price p):

$$\pi_R(p) = \frac{a-bp}{\bar{x}}\mathbf{E}\left[p\min\{Y, \bar{x}\} - w\min\{Y, x_0\}\right].$$

Again, according to Kouvelis et al. (2018), the unique optimal price can be solved from the FOC efficiently. Hence, we use bisect search over the interval $p \in [w, a/b]$ to find the optimal price p^*.

For the supplier, the profit function is

$$\pi_S = \mathbf{E}\left[w\hat{q} - cQ\right] = \frac{a - bp^*}{\bar{x}}\left(\hat{w} - c\right).$$

Now, given a, b, c, we can find the equilibrium outcome in the following procedure:

(1) For $w \in (c/\mu, a/b)$, use FOC regarding the function $\pi_R(p)$ to find optimal price $p^*(w)$; then we also have $q^*(w)$.

(2) Search $\pi_S(w)$ described above to find the optimal wholesale price w^*.

Scenario DR: Recall that the optimal price and revenue of the retailer is given by (4) and (8), respectively. In addition, the supplier will inflate production level in this scenario, and $Q = q/x_0$. Then, the retailer's profit can be written as a function of the delivered quantity $\hat{q}$; so the expectation in turn is a function of the order quantity. Specifically,

$$\pi_R(q) = \mathbf{E}\left[\frac{(a-\hat{q})\hat{q}}{b}\mathbf{1}_{\{\hat{q}<a/2\}} + \frac{a^2}{4b}\mathbf{1}_{\{\hat{q}\geq a/2\}} - w\hat{q}\right].$$

Clearly, the delivery $\hat{q} = \frac{q}{x_0}\hat{Y}$. Similar to the previous scenario, we consider two pieces. First, if $0 < q < a/2$, then $\pi_R(q) = \left(\frac{a}{b} - w\right)\mathbf{E}\hat{q} - \frac{1}{b}\mathbf{E}\hat{q}^2$, where $\hat{q} = q\hat{Y}/x_0$ and $\hat{Y} = \min\{Y, x_0\}$. Hence, the retailer's profit is a quadratic function. Second, if $a/2 \leq q < ax_0/(2x)$, then $\pi_R(q)$ is concave and the FOC, which leads to optimality (Kouvelis et al. 2018), is

$$\int_x^{\frac{ax_0}{2q}} \frac{1}{b}\left(ay - \frac{2y^2q}{x_0}\right) dF(y) = \hat{w}.$$

Recall $\hat{w} = w\left(\mu - \frac{(1-x_0)^2}{2(1-x)}\right)$. Therefore, given a w, the optimal order quantity of the retailer, $q^*(w)$, can be computed efficiently (by considering the two pieces and comparing their optima). Note that we do not need to consider the case of larger q because the profit function is decreasing on that region.

For the supplier, the profit function is

$$\pi_S(w) = \mathbf{E}\left[w\hat{q} - cQ\right] = \frac{q^*(w)}{x_0}(\hat{w} - c).$$

Now, given a, b, c, we can find the equilibrium outcome in the following procedure:

(1) For every $w \in (c/\mu, a/b)$, find optimal price $q^*(w)$ of the retailer based on the above discussion.

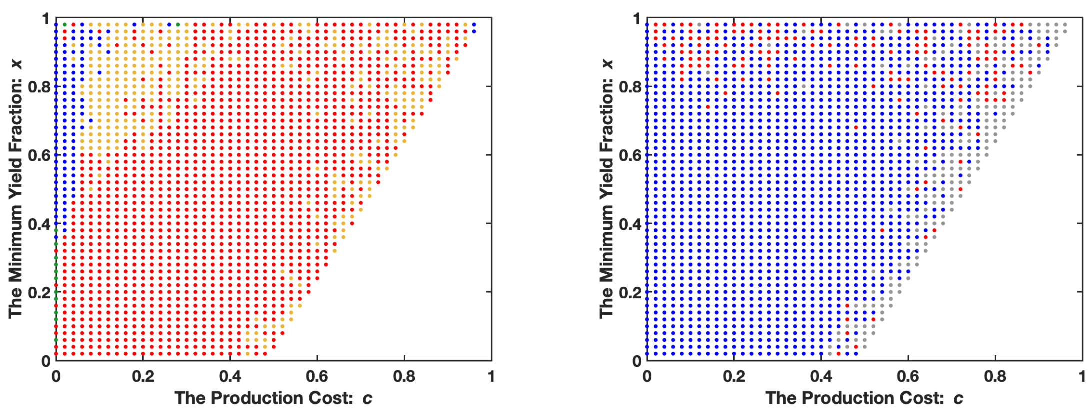
(2) Search to find the wholesale price w^* that maximizes the above profit function $\pi_S(w)$.

Design and Results Exhibitions of the Numerical Analysis: We mimic the studies in Section 6 by varying two parameters and using colored regions. Specifically, we fix the demand curve $a = b = 1$, change the minimum yield $x \in (0, 1]$ and the production cost $c \in [0, \mu]$, and color code the comparison results. It is worth mentioning that, although the x-axis is production cost in all figures here and in the main model, the y-axis is different for figures with different yield distributions. In the main model, the y-axis is θ, but here it is the minimum yield realization. Hence, the position of certain colored area in the corresponding figures may not be comparable. Below, we exhibit and discuss the results.

First, Figure 6(a) shows the retailer's optimal strategy and Figure 6(b) shows the interplay of the two postponement strategies. Figure 6(a) depicts the cases where the retailer prefers to stay status quo, especially when the production cost is small. Moreover, throughout the study, we did not observe the case that DA is optimal, with only few exception when $c = 0$. The intuition is as follows. When the yield is continuously distributed, the supplier always produces more than the order quantity and, as a result, charges higher wholesale price w in scenarios "D". Hence, the retailer must use responsive pricing to main the margin when being overcharged by the supplier.

Figure 6(b) echoes with the key observation from the main model. Notably, unlike Figure 2 in Section 6, in most of the region here the two strategies are complements. This is because, again, the supplier does not always inflate production for the two-point distribution case in the main model, whereas here the supplier will always share some supply risk whenever the retailer postpones the payment. The above discussions regarding Figure 6 support our remark (1) in Section 7.1.

Figure 6 **Strategy Selection and Interplay: Uniform Distributed Yield Fraction ($Y \sim U[x, 1]$)**



(a) Optimal Strategy for the Retailer.
Blue - OA; Red - DR; Yellow - OR; Green - DA.

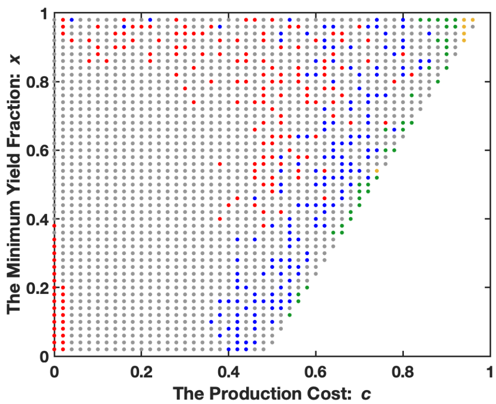
(b) The Interplay of the Two Strategies.
Blue - Comp.; Red - Sub.; Gray - Neither.

Second, Figure 7(a)&(b) show the impact of payment postponement on the entire channel. We can clearly observe that the channel impact of payment postponement can be positive under either pricing scheme. Moreover, similar to Figure 3 in Section 6.3, the blue area only takes a small portion of the whole region, meaning that the payment postponement is difficult in achieving the win-win-win equilibrium. Interestingly, we can observe the same finding here as in the main model: Under ex ante pricing, the consumers may be worse off (red region in Figure 7(a)) whereas under responsive pricing the supplier may be worse off (green region in Figure 7(b)).

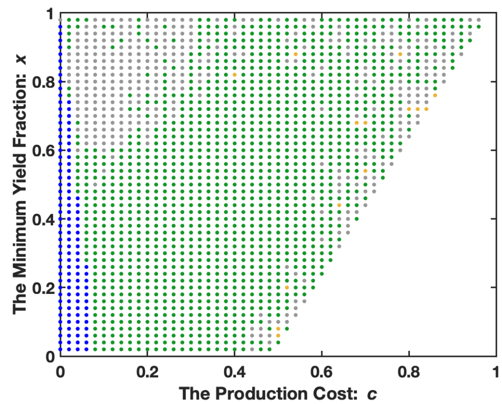
On the other hand, Figure 7(c)&(d) illustrate the channel impact of pricing postponement. Overall, the observation here is consistent with that for Figure 5 in Section 7.4. In particular, we can see that the pricing postponement can be Pareto optimal to the channel in most cases, especially when pay-by-delivery is used

– this is shown by Figure 7(d) where the majority of the figure is blue region (win-win-win case). Moreover, under pay-by-order contract, there are many cases where the consumers are hurt (red region in Figure 7(c)). The above discussions regarding Figure 7 support our remark (2) in Section 7.1.

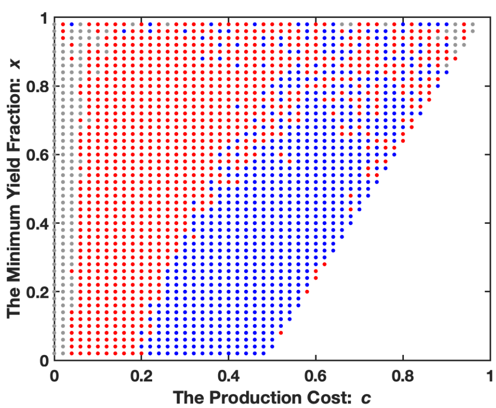
Figure 7 Channel Impact of Postponement Strategies: Uniform Distributed Yield Fraction ($Y \sim U[x,1]$)



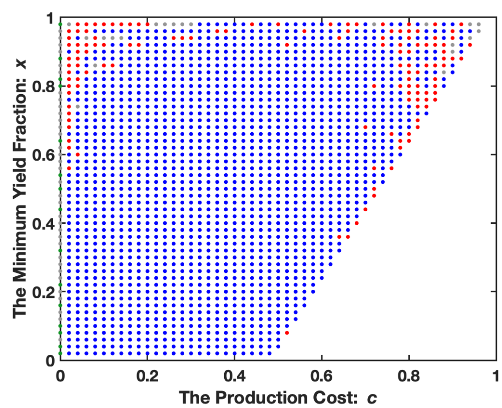
(a) Assume ex ante pricing.



(b) Assume responsive pricing.



(c) Assume a pay-by-order contract.



(d) Assume a pay-by-delivery contract.

Note: Blue - (W,W,W); Red - (W,W,L); Green - (W,L,W); Yellow - (W,L,L); Gray - (L,X,X).

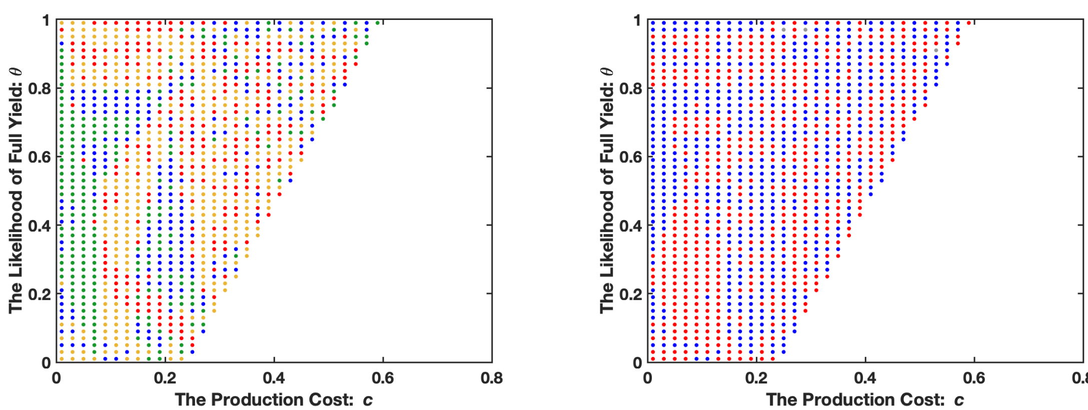
Lastly, we remark again that the uniform distribution is in nature dissimilar to the two-point distribution, and therefore the graphical results are not perfectly comparable. We only use these observations to show the qualitative robustness of our results. For example, despite the differences between Figure 7 and the results for the main model, the fact that postponement strategies can benefit the channel remains true in all cases.

Appendix E: Details of the Numerical Analysis for the Model with Demand Risk

We introduce the setup of our numerical analysis for the model incorporating the demand risk. In the numerical analysis, we look at the case where $\hat{a}$ follows a two-point distribution; i.e., the potential market size could either be high, $\hat{a} = a_H$ with probability γ, or low, $\hat{a} = a_L$ with probability $1-\gamma$. Specifically, we set $a_H = 1.2$, $a_L = 0.8$, and $\gamma := \mathbf{Prob}(\hat{a} = a_H) = 0.5$. In addition, the yield fraction is distributed in the same way as our main model, and we set $\alpha = 0.4$ (this is one of the numerical instances in Section 6). The slope in the demand function is fixed as $b = 1$. Lastly, recall that the upper bound for the production cost c in the main model is $a/b\mu$. In the presence of demand risk, the upper bound is squeezed by the probability of market size being large. Indeed, for the retailer, since the order is placed before the demand realizes, a necessary condition for its profit to be positive is $(p - w)\mathbf{Prob}(\hat{a} = a_H) - w\mathbf{Prob}(\hat{a} = a_L) > 0$. Hence, we have the production cost $c < w < p\mathbf{Prob}(\hat{a} = a_H) < a_H/b\mu\gamma$. This upper bound is applied throughout our numerical experiments. We follow the same procedure used in Section 6 to conduct a series of numerical studies and make several observations.

First, Figure 8(a)&(b) shows the retailer's strategy selection results. Specifically, (a) is the optimal strategy for retailer in different instances and (b) depicts the interplay of the two postponement strategies. These graphical illustrations confirm that our high-level takeaways in the main model are qualitatively unchanged when demand risk is considered. However, the region displays are more "mixed" than in the main model. This is due to the retailer making decisions facing a random demand (i.e., a result of coping with the demand risk). The expectation taken with respect to $\hat{a}$ has a considerable impact on the retailer's equilibrium profit. This observation leads to our remark (1) in Section 7.2.

Figure 8 Strategy Selection and Interplay: Incorporating Demand Risk ($a_H = 1.2$, $a_L = 0.8$, and $\gamma = 0.5$; $\alpha = 0.4$ and $b = 1$)

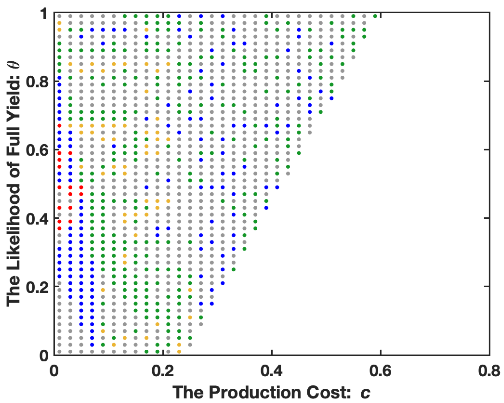


(a) Optimal Strategy for the Retailer.
Blue - OA; Red - DR; Yellow - OR; Green - DA.

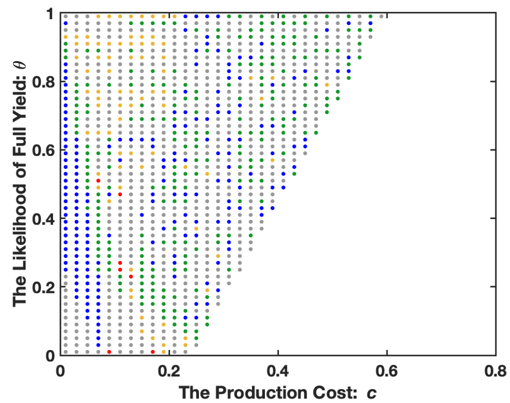
(b) The Interplay of the Two Strategies.
Blue - Comp.; Red - Sub.; Gray - Neither.

Second, Figure 9(a)&(b) show the impact of payment postponement on the entire channel. We can clearly observe that the channel impact of payment postponement can be positive under either pricing scheme. Moreover, similar to results shown in Section 6.3, it is not easy for the payment postponement to lead to a win-win-win situation. An interesting observation is that in Figures 9(a)&(b) there are both red region and green region simultaneously, which is in contrast to Figure 3 in Section 6.3 where consumers may be worse off *only* under ex ante pricing whereas the supplier may be worse off *only* under responsive pricing. On the other hand, Figure 9(c)&(d) illustrate the channel impact of pricing postponement. Overall, the observation here is consistent with that for Figure 4 in Section 6.4. In particular, the pricing postponement can be Pareto optimal (win-win-win situation) to the channel in most cases. However, here we have many other cases besides (W,W,W) where either consumers or the supplier, or both, may be hurt by the retailer's use of responsive pricing. Combining the above observations leads to our remarks (2) and (3) in Section 7.2.

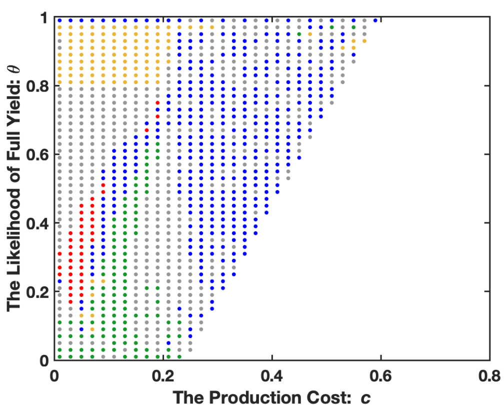
Figure 9 Channel Impact of Postponement Strategies: Incorporating Demand Risk ($a_H = 1.2$, $a_L = 0.8$, and $\gamma = 0.5$; $\alpha = 0.4$ and $b = 1$)



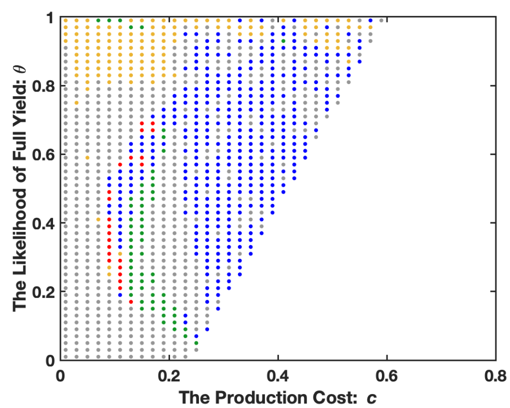
(a) Assume ex ante pricing.



(b) Assume responsive pricing.



(c) Assume a pay-by-order contract.



(d) Assume a pay-by-delivery contract.

Note: Blue - (W,W,W); Red - (W,W,L); Green - (W,L,W); Yellow - (W,L,L); Gray - (L,X,X).