POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT

Vol. 26, No. 8, August 2017, pp. 1553–1567
ISSN 1059-1478 | EISSN 1937-5956 | 17 | 2608 | 1553

DOI 10.1111/poms.12703

Using Advance Purchase Discount Contracts under Uncertain Information Acquisition Cost

Wenjie Tang
NUS Business School and Institute of Operations Research and Analytics, Singapore City 119245, Singapore, wenjie.tang@nus.edu.sg

Karan Girotra
Technology and Operations Management Area, INSEAD, Fontainebleau 77305, France, karan@girotra.com

We study the use of advance purchase discount (APD) contracts to incentivize a retailer to share demand information with a dual-sourcing wholesaler. We analyze such contracts in terms of two practical considerations that are relevant in this context but have been overlooked by previous work that has largely studied the direct offer of APD to customers: the retailer's information acquisition cost and the wholesaler's limited information about that cost. The wholesaler's limited knowledge of the retailer's cost leads to a departure—from the normal "full observability" APD design—that is asymmetric and depends on the extent of unobservability; if the uncertainty is small (resp., large) then the optimal discount is higher (resp., lower) than in the case of full observability. An APD contract that ignores the retailer's cost or the wholesaler's uncertainty about it will yield fewer benefits for the wholesaler and the supply chain. We offer a numerical illustration (calibrated on real industry data) establishing that for a representative product, an APD contract can improve the wholesaler's profit margin by as much as 3.5%.

Key words: supply chain management; information sharing; advance selling; demand forecasting
History: Received: May 2014; Accepted: February 2017 by Albert Ha, after 3 revisions.

1. Introduction

Supply–demand mismatches of products with short life cycle, e.g., apparel, electronics, vaccines, etc., are a central issue in operations management. Typically, the wholesaler of such products outsources the production overseas, accompanied by a much smaller scale local production with a higher production cost and shorter lead time. Wholesalers need to make ordering decisions for the overseas production well before the realization of the actual demand, with only very limited demand-relevant information at hand.

Retailers, on the other hand, often have access to better demand information than do upstream agents in the supply chain, thanks to their direct relationship with customers. One promising solution in these contexts is for the wholesaler to develop mechanisms for acquiring information from the retailer. Operations academics have suggested a variety of contracting schemes designed to achieve honest information sharing within a supply chain, including return and rebate contracts (Taylor and Xiao 2009), contracts with commitment and options (Cachon and Lariviere 2001, Özer and Wei 2006), and *advance purchase discount* (APD) contracts (Cachon 2004, Donohue 2000, Özer and Wei 2006, Özer et al. 2007). Among them, a typical APD contract should in theory incentivize the retailer to place an order well in advance of the peak sales period, allowing the wholesaler to infer the retailer's private information about demand from this early order quantity and thereby improve his own sourcing and production decisions. (To facilitate the exposition we use masculine and feminine pronouns for the wholesaler and retailer, respectively.)

Although new technological advances (such as internet websites, electronic cards, smart cards, etc.) have made it easier to implement APDs between firms and end customers (Boyaci and Özer 2010, McCardle et al. 2004, Shugan and Xie 2004), such customer-specific technologies have had little impact on the relationship between the upstream wholesaler and the downstream retailer, where the use of APD contracts is rarely documented. In such cases, a simple wholesale price contract (Cachon and Lariviere 2005, Kalkanci et al. 2011, Lovejoy 2010) or more specialized contracts (Dai et al. 2016) are often preferred.

Compared with advance selling contracts between the retailer and end customers, APD contracts between the wholesaler and the retailer are often more difficult to implement, and these challenges arise mainly due to the difference in the offer recipient in these two cases—the end customers receive the offer in the former case, whereas in the latter case the retailer receives the offer. Unlike the end customers (who are the essence of demand), the retailer is seldom endowed with demand information and must incur additional costs to obtain it. These costs are reflected not only in gathering the raw data (e.g., through conducting preseason merchandise tests and customer surveys; purchasing relevant sales or consumer data from third parties; installing item tracking and decision support systems such as barcodes, electronic data interchange, and radio frequency identification; and/or implementing various customer loyalty programs), but also in the time and resources devoted to analyzing data, consulting internal or external experts, and making relevant decisions (Aiyer and Ledesma 2004, Guo 2009, Hays 2004, Taylor and Xiao 2009). Furthermore, unlike direct customers who can be viewed as a large crowd and are often modeled as a population of infinitesimal agents that can be reasonably well characterized by a demand function, each retailer acts strategically, and thus has her own objective function that determines whether she should acquire information and participate in the proposed APD scheme. Therefore, if the discount is not properly tailored, the retailer might not have enough incentive to acquire information—owing, for example, to the additional costs that the retailer must incur to obtain demand-relevant information from the end customers. This discrete nature also complicates the upstream wholesaler's decision about how large a discount to offer. Last but not least, in practice these costs will depend on the retailer's internal mechanisms, organizational inertia, and other costs that are not visible to the wholesaler (Corbett 2001, Corbett and de Groote 2000, Corbett et al. 2004, Gurnani and Tang 1999, Ha 2001). These uncertainties pose additional challenges to the wholesaler, and make it even more difficult to successfully implement the APD scheme.

We propose an APD scheme that takes these considerations into account. In contrast to existing research, our design controls for the retailer's additional efforts to gather and process information and assumes that this cost is unobservable to the wholesaler. Furthermore, we explicitly model the retailer's decision whether or not to acquire information and participate in the APD scheme. We offer the following practical guidelines for a wholesaler setting an advance purchase discount: (i) First of all, the retailer's cost needs to be taken into account—an arbitrarily small discount is not acceptable. (ii) Moreover, the size (e.g., mean, range, first-order dominance) of the retailer's cost estimate alone is not adequate to determine an optimal discount—it is possible that a higher estimate leads to a lower discount due to the uncertainty about it. (iii) To cope with the uncertainty about the cost estimate, when the wholesaler is moderately uncertain about the retailer's cost, a higher discount should be given to increase the chances of the retailer's information acquisition; whereas when the uncertainty is fairly large, a lower discount is optimal.

Our results indicate that if the APD contract design ignores the retailer's information acquisition cost or the wholesaler's uncertainty about it, fewer benefits will accrue to the supply chain agents than would otherwise be the case. This finding may explain the limited use of APD contracts in practice, and provide new insights for better implementation of the APD scheme. Specifically, a wholesaler who does not account for the retailer's information acquisition cost will likely offer a discount that is insufficient to induce the retailer to acquire information. Furthermore, ignoring the uncertainty about the retailer's cost leads the wholesaler to offer a discount that could either be too low (in which case the retailer has no incentive to acquire information) or too high (in which case the benefit to the wholesaler is suboptimal).

To illustrate the APD implementation and to estimate the potential benefits of following our prescriptions, we provide a numerical study and estimate the revised APD scheme's benefits for the wholesaler and discuss practical issues associated with its implementation.

2. Motivating Examples

Our work is motivated by several examples across different industries. In the apparel industry, Costume Gallery is a New Jersey–based dance costume wholesaler that sells through dance schools to students and employs a combination of cheaper China-based sourcing and more expensive local production (Girotra and Tang 2009, 2010). The primary selling season for its products is in late April, and to meet this demand Costume Gallery must place its offshore production order in early February with very limited knowledge of what the demand might be for each costume. Hence in any given year, Costume Gallery is unable to sell as much as 35% of its stocked inventory and must produce as much as 20% of its sales using the more expensive in-house production resources. On the other hand, by the time Costume Gallery places its overseas production orders, a dance school already has some information that is useful for estimating the future demand for each costume SKU—such as what the dance theme would be in the late April performance and how many students in total are currently enrolled. This important information can potentially be conveyed to Costume Gallery if the dance school can place its order before Costume Gallery's China production order. However, that would not happen in the normal course of business. First of all, even for the dance school, there still remains considerable uncertainty on the final demand, mostly resulting from factors such as unfinished choreography, design, and allocation of roles, potential changes on the dance theme, students' costume sizes, and students' role assignment. Furthermore, dance school teachers must incur costs to obtain this information. For example, they need to finalize the choreography, design, and role assignments as well as decide on the best costume for each role in a time frame much tighter than before (when no such information acquisition was needed). The teachers must also measure the size of each student performer. In addition, the remaining uncertainties also make it challenging to estimate the future demand for different styles and sizes of costumes. This task requires careful thinking and deliberation that is based on currently available information but acknowledges the possibility of future deviations and their likely effects. Therefore, without specific incentives, dance school teachers would prefer waiting to place their orders only in April.

Managing the production of flu vaccines has similar challenges. Vaccines are known to have short life cycles since the prevalence of virus strains changes every year. The traditional egg-based production method is relatively cheaper and can take a long time, and hence manufacturers must start planning for production far in advance of the selling season (Chick et al. 2008, Dai et al. 2016, Heinrich 2001, O'Mara et al. 2003) with very little idea of what the future demand will be (Williams 2005). At the same time, manufacturers also employ faster and more costly cell-based production to cater to sudden increases in demand such as in the event of a pandemic (Health Sciences Authority 2009). As a result, mismatches of demand and supply occur regularly, and manufacturers have been found to rely on flawed heuristics when making their production decisions (Cachon and Terwiesch 2006, Rudi and Drake 2014). On the other hand, retailers such as government bodies, clinics, hospitals, health departments, senior centers, and large grocery chains are often better informed than the manufacturers about local medical facilities, demographics, the medical history of potential vaccine receivers, state policies that determine which populations and how many people should be vaccinated, and the prevailing local view for or against vaccination. This information can potentially be conveyed to the vaccine manufacturers if the retailers can place their order before the start of the production. However, the retailers do not have incentives to do so—aside from having to bear demand risk themselves in this case, to collect this information, retailers also need to incur costs. Therefore, in order to motivate retailers to place early orders and thereby to benefit from precious early demand information, vaccine manufacturers need to design proper incentives, by taking into account retailers' uncertain information acquisition cost.

In the electronics industry, a similar problem was faced by Xiaomi, a Chinese electronics company, when it first entered the Indian market in July 2014 and sold smartphones through the online megastore Flipkart. Xiaomi's primary production is in China, where the production is cheaper but takes a longer time (due to its large scale, further distance from India, and international shipping constraints). Concurrently, Xiaomi also produces locally at a much smaller scale and higher cost to cater to last-minute orders. Although familiar with its native Chinese market, Xiaomi knew little about India and had very little idea about demand when making decisions about its production in China. On the other hand, Flipkart in India knew better than Xiaomi what appearance, capacity, and functionality of smartphones would appeal most to local consumers. Such information could potentially be conveyed to Xiaomi if Flipkart can place its order before the start of the production. However, Flipkart also faces demand uncertainties and incurs costs from acquiring and consolidating this information, and therefore it has no incentive to collect this information or placing an early order.

The above three examples across the apparel, vaccine, and electronics industries all point to the useful private demand information accessible to the retailer, and the retailer's cost to obtain it. On top of that, since the wholesaler would know little about the cost structure within the organization of the retailer, he has little idea about the exact cost incurred by the retailer in order to obtain this information. These two factors constitute barriers to effectively implementing an APD scheme between a wholesaler and a retailer. Without taking them into account, retailers cannot be properly incentivized to place an early order, and therefore fail to convey the useful demand information—through the early order—to the wholesaler.

3. Related Literature

This study brings together and augments the literature on information sharing and advance purchase discount contracts.

3.1. Literature on Information Sharing Contracts

The pioneering work in this domain includes the well-known case of Sport Obermeyer (Fisher and Raman 1996, Fisher et al. 1994, Hammond and Raman 1994). Note that in the papers cited, the downstream agent's superior demand information is presented to the upstream agent as a matter of course. Our model builds on the basic setup described in the Sport Obermeyer case. However, instead of studying the value of information sharing, we focus on how to facilitate information sharing from the downstream to the upstream agent; this is a key issue in many business settings, including those in our motivating examples.

Previous work has identified different types of contracts that achieve the same goal in other contexts; these include return and rebate policies (Taylor and Xiao 2009), contracts with commitment and options (Cachon and Lariviere 2001, Özer and Wei 2006), and wholesale price contracts under confidentiality stipulations (Li and Zhang 2008).

In terms of supply chain structure, our proposed APD contract is designed to encourage information sharing between a dual-sourcing wholesaler and its retailer. In particular, unlike the stylized economic model of Li and Zhang (2008), which abstracts from the supply chain's internal physical flows, we model individual agents in the supply chain as being able to carry physical inventories and as facing the risk of over- or undersupply. Unlike the single-source capacity reservation model of Özer and Wei (2006), our study applies to settings where insufficient production capacity does not play a major role. And in contrast to Cachon and Lariviere (2001), in our model the upstream agent who needs the demand information (not the downstream one) offers the contract. As for the specifics of contract design, our proposed APD contract differs from that of Özer and Wei (2006)—who also consider a time-sensitive discount contract in a context of restricted production capacity—by endogenizing a previously exogenous contract parameter; namely, we identify the precise discount (in the APD contract) that optimally incentivizes a *self-enforcing* policy of information sharing. The proposed contract is simple and intuitive, since offering a discount for timely action often occurs in business, which makes it easier to introduce APDs within the constraints of an existing business culture.

Importantly, our model incorporates two practical concerns. First of all, it appropriately controls for the retailer's potential information acquisition cost. Furthermore, it explicitly considers a major practical constraint: the upstream agent's limited knowledge of the downstream agent's cost. We accommodate this limitation by offering actionable prescriptions and appropriate contract designs.

In short, the mechanism that we describe to facilitate information sharing differs from those proposed previously in terms of the applicable context, the simplicity and intuitiveness of the contract, and its robustness to realistic constraints.

3.2. Literature on Advance Purchase Discount Contracts

Advance purchase discounts or, more generally speaking, advance selling was first studied as direct offers from a retailer (or a manufacturer) to the end customers, aimed at obtaining advance demand information or implementing price discrimination policies (Boyaci and Özer 2010, Gundepudi et al. 2001, Li and Zhang 2013, McCardle et al. 2004, Prasad et al. 2011, Tang et al. 2004). More relevant to our study is the literature on APDs offered within a supply chain. The utilities and incentive structures of agents in a supply chain are usually assumed to differ from those of the end customers, leading to significantly different analyses and implications. In particular, supply chains involve issues of risk distribution and incentive coordination that are not relevant in the retailer–customer APDs. The pioneering works of Donohue (2000), Cachon (2004), and Özer et al. (2007) identified these concerns. However, neither Donohue (2000) nor Cachon (2004) addresses the information asymmetry between different supply chain tiers (i.e., upstream vs. downstream agents). Özer et al. (2007) do consider information asymmetry, but the downstream agent's private information is not shared within the supply chain. The recent work of Cho and Tang (2013) also investigates the option of advance selling in a supply chain that is characteristic of the flu vaccine industry, but the advance selling price is not necessarily lower than the regular one, and, more importantly, there is no information asymmetry between the two tiers.

4. Model Setup

In this section, we introduce a stylized model that captures the characteristics of our focal business setting and then discuss the implementation of an advance purchase discount contract.

4.1. Supply Chain Setup and the Information Environment

As in Figure 1, consider a wholesaler that sells a short life cycle product through a retailer. The retailer sources the product from the wholesaler with a nearly instantaneous lead time and then sells it in the market at a unit price p. We use D to denote the retailer's (selling season) demand.

Following Li and Zhang (2008), we assume that market demand D is normally distributed with mean μ_0 and variance $\sigma_0^2 = 1/\rho$, where ρ denotes the precision of this prior demand distribution. At some point into the natural process of business, private information about the market becomes accessible to the retailer. We model this private demand information as a signal Y of demand; we assume that $Y|D$ is distributed normally with mean D and finite variance $1/t$. The signal is thus an unbiased estimate of the true demand, where t can be interpreted analogously as the precision or quality of the retailer's market knowledge. We also assume that the retailer utilizes the information signal in a Bayesian fashion (Winkler

Figure 1 Setup of Supply Chain

1981). Hence, once the retailer, who has access to this private information signal, obtains it, she can use it to update the common prior on the demand distribution and thereby obtain the private posterior distribution of demand, $D \mid Y$, which is distributed normally with mean $\mu = (\rho\mu_0 + tY)/(\rho + t)$ and variance $\sigma^2 = 1/(\rho + t)$.

We assume that the retailer incurs cost k if she acquires the private demand information. It is a common assumption in the information-sharing literature that when the "informed" party has access to, but is not endowed with, the information, a fixed amount of cost will be incurred in order to obtain the information (Daughety and Reinganum 1994, Gabaix et al. 2006, Kurtulus et al. 2012, Taylor and Xiao 2009). Such costs reflect the time and attention required of the retailer to gather relevant information, conduct analysis, consult internal or external experts, and make relevant decisions (Aiyer and Ledesma 2004, Guo 2009, Hays 2004, Taylor and Xiao 2009). We assume throughout that the wholesaler is uncertain about the retailer's information acquisition cost (Corbett 2001, Corbett and de Groote 2000, Corbett et al. 2004, Gurnani and Tang 1999, Ha 2001) and in section 5.2 we specify how this uncertainty is incorporated into the model.

The wholesaler can potentially acquire these products from different sources, each with a different delivery lead time and associated costs. To simplify the analysis, we consider two extreme alternative sources for the product: it can be sourced (a) at unit cost c from a *local* source with a nearly instantaneous lead time, or (b) at unit cost $(1 - \gamma)c$, where $\gamma \in (0, 1)$, from an *offshore* source with a significantly longer lead time. In particular, any order from the offshore source must be placed well ahead of the product's selling season and so entails a make-to-*stock* inventory strategy. In contrast, a make-to-*order* strategy may be employed with respect to the local source. We assume that all the parameters described here, except for the private information signal Y, are common knowledge to both the wholesaler and the retailer.

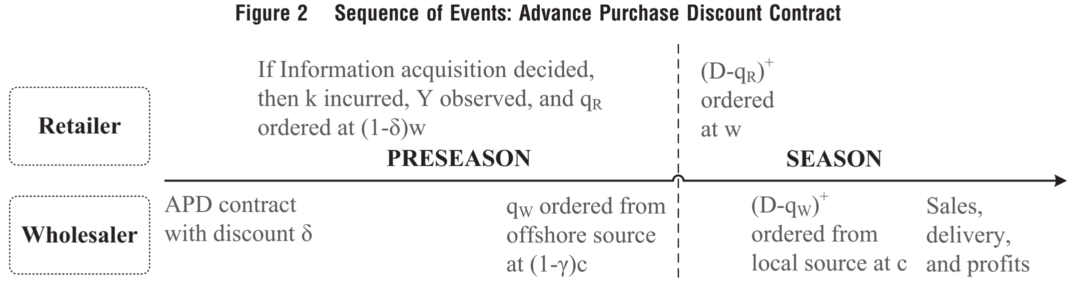
4.2. Advance Purchase Discount Scheme

We propose an APD contract whereby the wholesaler offers the retailer an opportunity to place an order in advance of the selling season at a discounted price of $(1 - \delta)w$ for $\delta \in [0, 1]$. The retailer may choose to acquire the private demand information Y at cost k, participate in the APD scheme, and purchase a certain quantity q_R during the preseason at this discounted price.

Alternatively, the two parties may consent to an exogenously given, conventional wholesale price contract. In this case, the wholesaler does not offer the retailer an opportunity to place an advance order, and the retailer orders an amount equal to the realized demand. We use the wholesale price contract as a benchmark: the wholesaler (resp., retailer) compares the outcomes from implementing the wholesale price contract and the advance purchase discount contract, and then decides whether to offer (resp., acquire information under) the APD contract.

The sequence of events under the APD contract is as illustrated in Figure 2. In advance of the selling

Figure 2 Sequence of Events: Advance Purchase Discount Contract



season, the wholesaler proposes the APD contract with discount δ. The retailer decides whether to acquire information or not. If the retailer does so, she incurs cost k to gather the private demand information Y that is accessible to her. She then uses that information to update the common prior on the demand distribution (note that once the retailer acquires information, she will always be better off using her private information about demand), thus obtaining the posterior distribution of demand $D \mid Y$; and orders q_R units at price $(1 - \delta)w$. On the other hand, if the retailer decides against it, she orders q_{R0} at the discounted price. After observing the retailer's order, the wholesaler places his order q_W from the offshore source at a unit cost of $(1 - \gamma)c$. Once demand D is realized, the retailer places an additional order of $(D - q_R)^+$ units at price w. The wholesaler then places orders $(D - q_W)^+$ from the local source when the offshore orders cannot satisfy all the realized demand. Finally, the delivery and payments are made and the retailer sells the product at unit retail price p.

5. Analysis

In this section we analyze the model introduced in section 4. Our aims are to illustrate the optimal strategies for—and benefits of—employing the advance purchase discount scheme (sections 5.1–5.3) and to examine the effect of the two practical concerns overlooked by previous research (section 5.4).

5.1. Retailer's Strategies

Under the APD contract, the retailer's strategy during the selling season follows directly from the setup described in section 4: the retailer simply orders from the wholesaler any quantities that are still needed to meet her realized demand. The preseason stage of the game is more interesting. Given a discount δ, if the retailer acquires information then her objective function is

$$\max_{q_R} \mathbb{E}_{D'}\big(pD' - w(1-\delta)q_R - w(D' - q_R)^+ - k\big), \quad (1)$$

where q_R denotes the retailer's order quantity and $D' = D|Y$ denotes the demand after observing signal Y. On the other hand, if the retailer decides against information acquisition, she does not update her demand information nor incurs any cost, and her objective function becomes $\max_{q_{R0}} \mathbb{E}_D(pD - w(1-\delta)q_{R0} - w(D - q_{R0})^+)$. The following lemma characterizes the retailer's strategy in response to the wholesaler's offer. To avoid trivial cases where a retailer never finds it appealing to acquire information due to her high information acquisition cost, we limit the retailer's cost to $k \le \hat{k}$, where $\hat{k} = w\phi(0) \cdot (\sigma_0 - \sigma)$. (All proofs are given in the Appendix to this paper).

Lemma 1. *A retailer always prefers the APD contract to a wholesale price contract. Furthermore, when the discount offered by the wholesaler is neither too small nor too large, that is, when $\underline{\delta} \le \delta \le \bar{\delta}$), the retailer acquires information and orders quantity q_R; otherwise, the retailer does not acquire information and orders q_{R0}. Here*

$$\begin{aligned} q_R &= \mu + \sigma z_\delta, q_{R0} = \mu_0 + \sigma_0 z_\delta, \text{ and} \\ \underline{\delta}, \bar{\delta} &= \{\delta | w\phi(z_\delta)(\sigma_0 - \sigma) = k\}, \end{aligned} \quad (2)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the pdf and cdf of the standard normal distribution, $z_\delta = \Phi^{-1}(\delta)$, and $0 < \underline{\delta} < \frac{1}{2} < \bar{\delta} < 1$.

The retailer's decision consists of two steps: whether to acquire information and how much to order. We demonstrate the existence of a range of discounts $[\underline{\delta}, \bar{\delta}]$ such that the retailer acquires information only if the discount offered by the wholesaler is within the range; otherwise, the option of not investing in demand forecasting is more appealing to the retailer. We refer to the lower bound $\underline{\delta}$ as the retailer's minimum acceptable discount (MAD). Not surprisingly, the threshold level is increasing in k; thus, as the retailer's cost of acquiring information increases, a deeper discount is required.

Recall that if the retailer decides to acquire information, then she chooses her order quantity on the basis of all available information. In particular, she uses her private information about demand to update the common prior and then uses the resulting posterior estimate of demand. This leads to an expression for q_R in Equation (2) that is a function of the posterior estimate. Recall that μ, the posterior estimate of the mean of the demand distribution, is a function of Y; hence the retailer's order quantity q_R is a fully invertible function of her private signal Y. So by observing the retailer's order, the wholesaler can accurately infer that private signal.

5.2. Wholesaler's Strategies

The wholesaler's preseason strategy also consists of two key decisions: what discount δ to offer, and how many units q_W to order from the offshore source. These decisions must be made while bearing in mind the retailer's reaction to the offered discount. Assuming that the retailer acquires information with discount δ and the retailer's early order quantity q_R, the wholesaler's objective function is

$$\begin{aligned} \max_{q_W} \mathbb{E}_{D'}\big(&w(1-\delta)q_R + w(D' - q_R)^+ - c(1-\gamma)q_W \\ &- c(D' - q_W)^+\big). \end{aligned}$$

The next lemma characterizes the wholesaler's optimal order quantities and the resulting benefits from the implementation of the APD scheme.

LEMMA 2. *Given discount δ and that the retailer acquires information, the wholesaler would order $q_W = \mu + \sigma z_\gamma$ with resulting expected benefits from the APD scheme*

$$\Delta_W = -w\Phi(z_\delta)\mu_0 + w\sigma\phi(z_\delta) + \bar{k}, \quad \text{where } \bar{k} = c\phi(z_\gamma)(\sigma_0 - \sigma). \tag{3}$$

Several effects compose the wholesaler's expected benefits: (a) By availing themselves of the APD, the average transfer price between the wholesaler and the retailer is reduced; this effect is captured by the term $w\Phi(z_\delta)\mu_0$. (b) In availing herself of the APD, the retailer signals to the wholesaler her private information about demand, which the wholesaler can utilize to update his demand forecast and reduce its variance from σ_0^2 to σ^2. This allows the wholesaler to better match his overseas order with the market demand and thereby reduce the demand–supply mismatch costs, as captured by the term $\bar{k} = c\phi(z_\gamma)(\sigma_0 - \sigma)$. Therefore, here $\bar{k}$ can be considered as the value of information to the supply chain. (c) Finally, in availing himself of the APD, the wholesaler ends up transferring some of the oversupply risk to the retailer, reducing his oversupply risk by the same amount. This is captured by the term $w\phi(z_\delta)\sigma$. A similar risk-transfer effect is captured by the model in Cachon (2004), but its lack of information asymmetry means that there is no interplay with the information effect.

Note that whereas a better informed retailer will be more beneficial to the wholesaler in (b) getting better demand information, it serves the opposite in (c) the risk-sharing aspect—a worse informed retailer can share more supply–demand mismatch risk with the wholesaler. Therefore, whether a better- or worse-informed retailer will be preferred by the wholesaler depends on how important the two aspects—information and risk sharing—are, respectively, to the wholesaler. The importance or weight of these two aspects is captured by $w\phi(z_\delta)$ and $c\phi(z_\gamma)$. We will show in section 5.3 that once we internalize the wholesaler's optimal discount decision, we will always have that $w\phi(z_{\delta^*}) < c\phi(z_\gamma)$, that is, as long as it is optimal for the wholesaler to offer the APD scheme to the retailer, he values information more than risk sharing.

To optimize such benefits, the wholesaler faces a trade-off between the likely results of a deeper discount and a more modest one. A modest discount yields more benefits to the wholesaler from the APD scheme but reduces the likelihood of the retailer acquiring information; a deep discount increases the chances of information acquisition, but at the cost of reduced benefits for the wholesaler. Intuitively, if the wholesaler can fully observe the retailer's cost, then he is able to calculate the retailer's MAD and set the discount just high enough to induce her to acquire information; he then orders based on the information inferred from the retailer's advance order quantity. However, his uncertainty about k complicates the decision.

To characterize the wholesaler's uncertainty about the retailer's cost k, we assume that in the wholesaler's subjective belief, k follows a distribution with a general pdf $g(\cdot)$, a cdf $G(\cdot)$, and positive support in the interval $[A, B]$. We also assume that the pdf of the distribution is log-concave; this assumption is not restrictive because it includes most common distributions such as the Weibull, gamma, normal, truncated normal, truncated logistic, uniform, and many other discrete distributions (Rosling 2002). Under such assumptions, the wholesaler's expected benefit from the APD scheme becomes

$$\begin{aligned}\mathbb{E}_k\Delta_W(\delta) = &Prob(w\phi(z_\delta)(\sigma_0 - \sigma) - k \geq 0)(-\Phi(z_\delta)w\mu_0 \\ &+ w\phi(z_\delta)\sigma + c\phi(z_\gamma)(\sigma_0 - \sigma)) \\ &+ Prob(w\phi(z_\delta)(\sigma_0 - \sigma) - k < 0)(-\Phi(z_\delta) \\ &\cdot w\mu_0 + w\phi(z_\delta)\sigma_0)\end{aligned} \tag{4}$$

where the first part denotes the wholesaler's expected benefit from the APD scheme when the discount δ is large enough to induce the retailer to acquire information; the second represents the case when it is not, such that the wholesaler endows the retailer with the discount benefit but is rewarded with only the risk-sharing benefit, which is not enough to justify the discount given.

Intuitively, the wholesaler's strategy can be characterized as follows. Let $\delta^* = \text{argmax}_{0<\delta\leq\frac{1}{2}}\mathbb{E}_k\Delta_W(\delta)$, where $\mathbb{E}_k\Delta_W(\delta)$ is defined as in Equation (4). If δ^* is such that $\mathbb{E}_k\Delta_W(\delta^*)$ is non-negative, then the wholesaler offers the APD scheme with discount δ^*: if the retailer does acquire information under such a discount, the wholesaler places an offshore order based on the updated belief about demand, D'; otherwise, the order is placed based on the prior demand distribution. On the other hand, if $\mathbb{E}_k\Delta_W(\delta^*)$ is negative, the wholesaler does not offer the APD scheme, and the wholesale price contract assumes.

In order to find the solution of $\delta^* = \text{argmax}_{0<\delta\leq\frac{1}{2}} \mathbb{E}_k\Delta_W(\delta)$, we first define $k^* = w\phi(z_{\delta^*})(\sigma_0 - \sigma)$. In general, $w\phi(z_\delta)(\sigma_0 - \sigma)$ is the value of information to the retailer for discount δ, and $w\phi(z_\delta)(\sigma_0 - \sigma) = k$ defines a one-to-one relationship between δ and k, for $0 < \delta \leq \frac{1}{2}$ and $k \leq \min(\bar{k}, \hat{k})$. Therefore, the optimization problem can be cast in terms of k, and once k^* is obtained, we can find the optimal discount δ^* through

$$\delta^* = \left\{\delta \mid w\phi(z_\delta)(\sigma_0 - \sigma) = k^*, \delta < \frac{1}{2}\right\}. \tag{5}$$

5.3. Equilibrium Outcome

The equilibrium outcome follows from the strategy profiles just described for the retailer and the wholesaler. Our next proposition formalizes this outcome.

PROPOSITION 1. *When $\mathbb{E}_k\Delta_W(\delta^*)$ is non-negative, the tuple of actions $\{\delta^*, q_R^*, q_W^*\}$ characterizes a perfect Bayesian equilibria of the setup just described, where*

$$q_R^* = \mathbf{1}_{k \le k^*}(\mu + \sigma z_{\delta^*}) + \mathbf{1}_{k > k^*}(\mu_0 + \sigma_0 z_{\delta^*}) \text{ and}$$
$$q_W^* = \mathbf{1}_{k \le k^*}(\mu + \sigma z_{\gamma}) + \mathbf{1}_{k > k^*}(\mu_0 + \sigma_0 z_{\gamma})$$

The equilibrium benefits from the APD scheme for the retailer, the wholesaler, and the supply chain are, respectively,

$$\begin{aligned}\Delta_R^* =& \mathbf{1}_{k \le k^*}\left(w\Phi(z_{\delta(k^*)})\mu_0 - w\phi(z_{\delta^*})\sigma - k\right)\\ &+ \mathbf{1}_{k > k^*}(w\Phi(z_{\delta^*})\mu_0 - w\phi(z_{\delta^*})\sigma_0),\\ \Delta_W^* =& \mathbf{1}_{k \le k^*}\left(-w\Phi(z_{\delta^*})\mu_0 + w\phi(z_{\delta^*})\sigma + \bar{k}\right)\\ &+ \mathbf{1}_{k > k^*}(-w\Phi(z_{\delta^*})\mu_0 + w\phi(z_{\delta^*})\sigma_0),\\ \Delta_S^* =& \mathbf{1}_{k \le k^*}(\bar{k} - k) \ge 0.\end{aligned}$$

Here $z_{\delta^*} = \Phi^{-1}(\delta^*)$ and $z_\gamma = \Phi^{-1}(\gamma)$. On the other hand, if Equation (4) is negative under δ^*, the wholesaler does not offer the APD scheme, and the wholesale price contract assumes.

Since the retailer's cost k is unobservable, the wholesaler assumes that her true cost is k^* and offers an APD contract that reflects this estimate. This estimate, k^*, can be interpreted as the certainty equivalent of the retailer's uncertain cost. The retailer always shares supply–demand mismatch risk with the wholesaler in exchange for a discounted wholesale price; furthermore, if $k^* > k$, then the retailer acquires information, and improves her own demand forecast so as to share less risk with the wholesaler, at the cost of k. The wholesaler, on the other hand, always loses the discounted margin to the shared risk; if $k^* > k$ then he gains from better demand forecast and inevitably allocates less risk to the retailer. As a result, the supply chain positively benefits from the APD scheme when the wholesaler manages to incentivize the retailer's information-acquisition behavior.

Note that under the equilibrium, we also have $\delta^* \le \gamma$, that is, the optimal discount set by the wholesaler is always no larger than the discount that the wholesaler can get from the offshore production source. This is because that one necessary (though not sufficient) condition for k^* is that $k^* = w\phi(z_{\delta^*})(\sigma_0 - \sigma) \le \bar{k} = c\phi(z_\gamma)(\sigma_0 - \sigma)$, that is, it is optimal for the wholesaler to offer the discount to the retailer when he considers her information acquisition cost to be no higher than what the information is worth to the entire supply chain. And then $\delta^* \le \gamma$ follows from the fact that $w \ge c$ and $\delta^* < \frac{1}{2}$. This condition then naturally ensures that the wholesaler's early production order q_W^* will always be no smaller than the retailer's early order q_R^*.

5.4. Impact of Practical Concerns

In this section we analyze how an APD scheme's implementation is affected by the two considerations described previously: the retailer's information acquisition cost, and the wholesaler's uncertainty about that cost.

5.4.1. Information Acquisition Cost.

If the wholesaler has full knowledge of the retailer's cost k, then he can compute her MAD as in Equation (2) and set the discount just high enough to induce her to acquire information. Thus the optimal discount is strictly a function of k's magnitude. A lower k implies a lower optimal discount, and *vice versa*. As long as the retailer's cost is small enough, the wholesaler offers an APD with discount $\delta = \underline{\delta}$, such that the retailer's gain from information acquisition just offsets her cost k.

If the wholesaler fails to account for the retailer's cost, then the discount he offers will not be enough to induce her to acquire information, and hence the wholesaler will not benefit from the implementation of the APD scheme.

5.4.2. Wholesaler's Uncertainty about Retailer's Cost.

When the wholesaler cannot observe k, he treats it as a random variable with a probability distribution. In this case, the optimal discount will depend on not only the "size" of k—e.g., mean, range, or first-order dominance—but also other aspects of the probability distribution. This claim is illustrated by the following example.

Consider three retailers, each of whom has a cost k that is unobservable to the wholesaler. For simplicity, we suppose that this cost can be either low (k_l) or high (k_h) with equal probability. Table 1 shows $\mathbb{E}[k]$, k^*, and δ^* for each retailer.

The costs for both retailer 2 and retailer 3 are, on average, higher than those for retailer 1, and they are also larger in the sense of first-order stochastic dominance. In the absence of uncertainty considerations, Lemma 1 suggests that the discount offered to either

Table 1 Discounts for Three Retailers with Different Cost Distributions. The results are calculated while assuming c = 1.8, w = 1.9, γ = 0.5, μ_0 = 10, σ_0 = 3, and σ = 0.1

Retailer	k_l	k_h	$\mathbb{E}[k]$	k^*	$\delta^*(\times 10^{-3})$
1	0.1	0.2	0.15	0.2	14.28
2	0.1	0.3	0.2	0.3	22.98
3	0.1	2	1.05	0.1	6.46

retailer 2 or 3 would be higher than the discount offered to retailer 1. This logic holds for the case of retailers 1 and 2, but the optimal discount offered to retailer 3 is actually lower than that offered to either retailer 1 or 2! In particular, not knowing the exact cost of the retailers, the wholesaler treats retailer 1 and 2 as if their costs equal the corresponding k_l—which is greater than the mean, $\mathbb{E}[k]$—and sets the discount accordingly. On the other hand, it is optimal for the wholesaler to treat retailer 3 as if her cost equals the corresponding k_h—which is less than $\mathbb{E}[k]$. This example illustrates that incorporating the wholesaler's uncertainty about the retailer's cost alters the landscape of optimal discounts: neither the unknown cost's mean nor its first-order stochastic dominance alone determines the optimal discount. The following proposition defines a stricter criterion.

PROPOSITION 2. *Suppose that both $(\bar{k} - k)g_1(k) - G_1(k)$ and $(\bar{k} - k)g_2(k) - G_2(k)$, $k \leq \bar{k}$ are increasing in k, and that $(\bar{k} - k)g_1(k) - G_1(k) \leq (\bar{k} - k)g_2(k) - G_2(k)$. Then the optimal discount, δ^*, is no less when k is drawn from the distribution with pdf $g_2(\cdot)$ than when k is drawn from the distribution with pdf $g_1(\cdot)$.*

A larger $(\bar{k} - k)g(k) - G(k)$ implies first-order stochastic dominance. Proposition 2 shows that the wholesaler's optimal discount could be ranked by the $(\bar{k} - k)g(k) - G(k)$ of the retailer's cost. As the retailer becomes more difficult to please—that is, as the distribution of k improves in terms of $(\bar{k} - k)g(k) - G(k)$ (implying that a deeper discount is needed to incentivize her information acquisition)—the optimal discount level increases. Next we dig deeper to consider the effect of the wholesaler's uncertainty about k on his optimal offered discount: in particular, should the wholesaler offer a larger or smaller discount in the presence of uncertainty?

Write the retailer's cost as $k = u + s\varepsilon$; here ε is a random variable whose pdf $g_0(\varepsilon)$ is log-concave with standard deviation equal to 1. Then k has cdf $G(k|u, s) = G_0((k - u)/s)$ and a decreasing reversed hazard rate. The standard deviation of the distribution $G(k \mid u, s)$ is equal to s, and increasing (resp., decreasing) s corresponds to expanding (resp., contracting) the distribution of k around u. In this sense, s serves as a measure of the wholesaler's uncertainty about the retailer's cost k. Observe that when $s = 0$, the distribution of k is degenerate with the mass point at u, then the optimal discount to offer is $\delta^* = \delta(u)$. The next proposition summarizes the impact of uncertainty on the optimal discount.

PROPOSITION 3. *Denote by $k^*(u, s)$ the solution to Equation (A1) when k follows a distribution with cdf $G(k|u, s) = G_0((k - u)/s)$. The optimal discount is $\delta^* = \delta(k^*(u, s))$, where $\delta(\cdot)$ is as defined in Equation (5). Then (i) the optimal discount first increases then decreases in the uncertainty s. Furthermore, (ii) if $s \leq s_0$, then $\delta^* = \delta(k^*(u, s)) \geq \delta(u)$, and if $s > s_0$, then $\delta^* = \delta(k^*(u, s)) < \delta(u)$. Here $s_0 = (\bar{k} - u)g_0(0)/G_0(0)$.*

Proposition 3 offers the wholesaler an important practical guideline for setting the discount when dealing with a retailer about whom he has limited knowledge. The key to setting the appropriate discount lies not only in the estimate of the unobservable cost but also in the uncertainty around it. In essence, if the uncertainty about k is relatively small, then a higher discount should be offered to be more sure that the retailer acquires information. However, this increase is not monotone: when the uncertainty is large enough, it is actually optimal to offer a lower discount. This is because when k is uncertain, the wholesaler needs to choose between a deeper discount and a more modest one—a modest discount yields the wholesaler a higher margin but reduces the likelihood of the retailer acquiring information, whereas a deep discount increases the chances of information acquisition but at the cost of a reduced margin for the wholesaler. When the uncertainty is small, it is within the wholesaler's capacity to offer a higher discount to increase the chances that the retailer acquires information (e.g., in the first two cases of Example 1, the discount offered ensures that the retailer of either type acquires information); but when this becomes a far stretch, that is, the discount required to incentivize the retailer's information acquisition becomes too high, the wholesaler would rather give up certain information acquisition probability of the retailer to protect his margin (e.g., in the last case of Example 1, the discount is only sufficient to incentivize the low-cost retailer to acquire information). Armed merely with these two intuitive measures and the insights from our foregoing discussion, the wholesaler is well equipped with guidance on how to proceed when offering APDs to his supply chain partner.

Interestingly, from the retailer's perspective, this proposition explicates her strategic incentive to communicate information about her cost. She will receive the deepest discount when the wholesaler's uncertainty about her cost is neither too high nor too low. In order to encourage the wholesaler to offer a high discount, the retailer should not completely hide her cost information from the wholesaler; doing so would lead him to react with a safe strategy—namely, setting a modest discount to ensure that he secures most of the APD benefit. Yet neither should the retailer communicate her cost information fully, since then the wholesaler would take full advantage and set the discount equal to her minimum acceptable level (thus minimizing the retailer's APD benefit).

The literature on behavioral operations management has shown that contracts proposed by a wholesaler who has incomplete information about the retailer's cost are often rejected (Katok and Pavlov 2013, Katok et al. 2014). In line with this, our model shows that APD implementation is completely different in cases with vs. without the wholesaler's uncertainty about the cost. We conjecture that failing to account for this uncertainty correctly—and thus not offering a proper discount to induce the retailer to acquire information—is a major reason why APD contracts within the supply chain are seldom observed in practice.

Note that if the wholesaler neglects the uncertainty s, then, according to the discussion after Lemma 1, he would offer a discount of $\delta(u) = \{\delta \mid w\phi(z_\delta)(\sigma_0 - \sigma) = u\}$, where $z_\delta = \Phi^{-1}(\delta)$. The following corollary summarizes the impact that neglecting uncertainty has on APD benefits to the wholesaler, the retailer, and the supply chain.

COROLLARY 1. *When deciding how much of a discount to offer, if the wholesaler ignores his uncertainty about the retailer's cost, then he is always worse off in terms of expected benefit. Both the retailer and the supply chain are also worse off if that ignored uncertainty is large* (*i.e., when* $s > s_0$) *but better off otherwise* (*i.e., when* $s \leq s_0$).

According to Proposition 3, if the ignored uncertainty is high, then the wholesaler offers a larger-than-optimal discount (from his standpoint), which is unnecessarily lucrative for the retailer; hence both the retailer and the supply chain are better off in terms of expected benefits, albeit at the expense of the wholesaler. Yet if the ignored uncertainty is small, then both parties (and the supply chain) are worse off owing to the smaller-than-optimal discount. We emphasize that the successful implementation of an APD scheme crucially depends on the wholesaler taking this uncertainty into account. Our findings suggest that neglecting or misinterpreting uncertainty may explain why —despite their potential benefits—APD schemes are neither trivial to implement nor commonly observed in practice.

6. Numerical Experiments

Our theoretical model suggests that, by addressing the two practical concerns discussed in previous sections, a wholesaler can better implement his advance purchase discount scheme; he could then elicit accurate and timely information from a retailer that can be used to improve his production decisions. In this section, we provide a numerical study to illustrate our analytical findings and to discuss the implementation of the APD scheme in practice.

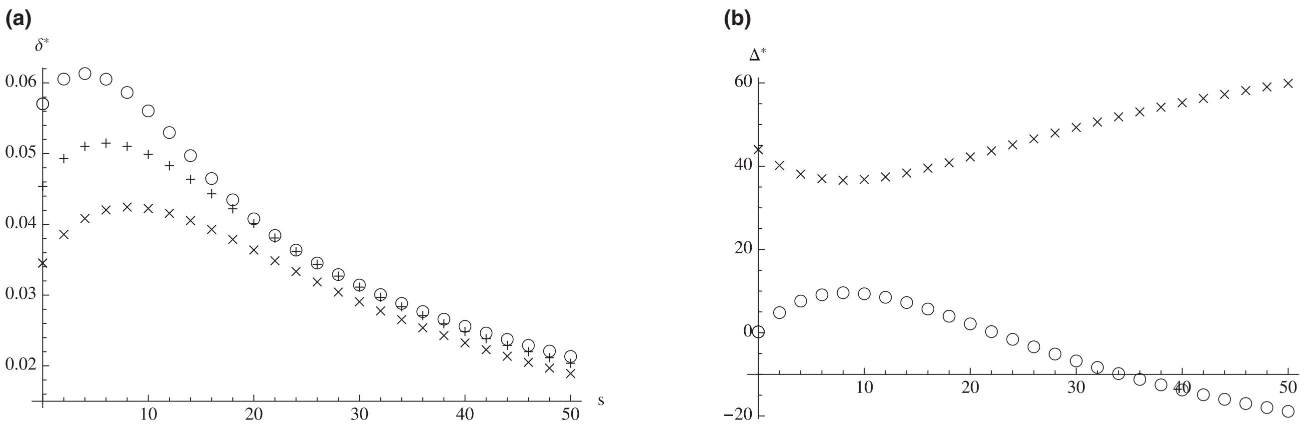
Assume that each product is offered to the retailer at a price of \$35 (w = 35 in our model), whereafter the retailer sells it to a customer at a price of \$40 (p = 40). It costs the wholesaler \$14 (c = 14) to produce it in-house, but if the wholesaler outsources the production, then the total cost—including shipping and all other incidentals—is estimated to be \$7 (γ = 0.5). The demand D from our model corresponds to the number of units requested by the retailer of a particular type and with a particular specification. This corresponds to demand for one stock-keeping unit to be sourced. We refer to this as an *order*, and we refer to D as the *order size*. Assume that the wholesaler's demand forecast is normally distributed, with a mean of 44.3 units and a standard deviation of 19.4 units ($\mu_0 = 44.3$, $\sigma_0 = 19.4$) and that errors in the wholesaler's own forecast can be reduced by about 75% when using advance information from the retailer. Then it follows that the posterior demand distribution has a standard deviation that is half that of the prior demand distribution—in other words, σ = 5.

Next, the wholesaler must estimate the retailer's information acquisition cost. This cost k depends on the retailer's internal mechanisms, organizational inertia, and other processes that are invisible to the wholesaler. The wholesaler cannot reliably estimate such a cost, which is believed to vary widely from one retailer to the next: some retailers are smaller in size and have a simpler organizational structure than others. The wholesaler also has a better sense of the cost at retailers with which he has long collaborated than at retailers that are new clients. To capture this heterogeneity in retailer types, we view any given retailer as having different estimated costs and allow for a range of uncertainty regarding those costs.

Figure 3a shows how the equilibrium discount varies with uncertainty about different estimated costs. For a given degree of uncertainty, higher estimated costs should correspond to a larger discount offered by the wholesaler to induce a retailer to acquire information in the APD scheme. In accordance with Proposition 3, however, the discount is not monotonic in the cost uncertainty. For instance, for a retailer with whom the wholesaler has a long-term relationship, the uncertainty is relatively small and so a deeper discount (than if the wholesaler knew the retailer's cost for certain) should be offered in order to increase the likelihood of that retailer's information acquisition. In contrast, it is safer for the wholesaler to offer a shallower discount to a new retailer, assuming the same estimated cost.

Next we examine benefits from implementing the APD scheme. For the purpose of illustration, we take the case where the cost estimation equals \$40, and we assume that the retailer's real cost equals the estimation, that is, k = 40. The next figure shows the benefits

Figure 3 Effect of Uncertainty Levels on Equilibrium Discounts and Benefits. (a) Equilibrium discount δ^* as a function of the uncertainty s surrounding estimated costs of u = 40 (×), u = 50 (+), and u = 60 (◯). (b) Equilibrium benefits Δ^* for the wholesaler (×) and the retailer (◯) as a function of the uncertainty s surrounding an estimated cost of u = 40



for the retailer as well as for the wholesaler over a range of uncertainty around the estimated cost.

Figure 3b shows that for the focal product with the given specification, it is always beneficial for the wholesaler to offer an APD scheme. Successfully implementing the APD scheme allows the wholesaler to increase its profit by \$40 (3.5%) on average.

7. Conclusion

This study addresses two practical considerations in information-acquiring contracts that are largely ignored in the existing literature. First, the downstream retailer is not endowed with but only has access to private information, hence to obtain it she must spend resources and thereby incur additional costs (Aiyer and Ledesma 2004, Guo 2009, Hays 2004, Taylor and Xiao 2009). Second, those costs depend on the retailer's internal mechanisms, organizational inertia, and other costs that are not visible to the wholesaler (Corbett 2001, Corbett and de Groote 2000, Corbett et al. 2004, Gurnani and Tang 1999, Ha 2001). When designing contracts, it is essential that the wholesaler accounts for this uncertainty about the retailer's costs.

We propose the mechanism of an advance purchase discount scheme that accommodates these considerations. We assume that the retailer incurs an information acquisition cost to obtain private demand information that is available to her, and that this cost is unobservable to the wholesaler. We characterize the retailer's optimal strategy as a "minimal acceptable discount" policy; thus she acquires private demand information only if the wholesaler's offered discount is no less than her threshold value. In light of that threshold, we then provide the optimal discount for the wholesaler. We find that the size of the optimal discount (from the wholesaler's standpoint) cannot solely be determined by the mean of the unknown cost nor that cost's first-order stochastic dominance. In particular, our analysis focuses on how the degree of uncertainty affects the wholesaler's optimal offered discount. We discover that, when uncertainty in the unobserved cost is low (resp., high), he should offer a deeper (resp., shallower) discount than in the case of full observability.

An important finding of this research is that a contract whose design ignores the retailer's information acquisition cost and the wholesaler's uncertainty about it is likely to yield less benefit to supply chain agents. This may explain the limited use of APD contracts in practice despite their demonstrated potential advantages. A wholesaler who does not account for the retailer's information acquisition cost would offer a relatively small discount, one that would be insufficient to induce her to acquire demand information. Moreover, if uncertainty about the cost is ignored, then the offered discount will likely be either too low or unnecessarily high. We argue that failing to accommodate either (*a fortiori* both) of these practical concerns will compromise the implementation of the APD scheme and thus constrain its use in practice.

Finally, we provide a numerical example to illustrate the implementation of the APD and to estimate its benefits for the wholesaler, the retailer, and the supply chain.

The APD proposal described here has several limitations. Designing a truly optimal APD contract would require estimation of retailer parameters to a degree of precision that is not practically feasible. Our analysis with limited knowledge about retailer cost addresses these concerns to some extent; however, a viable implementation of any APD scheme would require—in addition to precise cost accounting systems—that the wholesaler have extensive archival data on his own (and the retailer's) forecasting ability in order to accurately measure the benefit of offshore production. These conditions are not likely to be satisfied by many wholesalers. Alternatively, we can extend our current model to include uncertainty for other parameters of the retailer and the wholesaler, e.g., the retailer's signal precision. Note also that the scheme's benefits are highly sensitive to the extent of the offered discount, yet computing and administering that discount requires managerial skills that may not be available to small organizations. Moreover, practical constraints arising from industry practice and long-held relationships with supply chain partners could well prevent a wholesaler from implementing the optimal APD contract, which in turn would significantly reduce the accrued benefit. Finally, the proposed scheme is geared to a supply chain consisting of one upstream wholesaler and one downstream retailer. Yet wholesalers (such as Costume Gallery, Xiaomi, etc.) almost always deal with multiple retailers. In that case, a single discount will not adequately distinguish among retailers with varied costs, but since this private cost information is not associated with any retailer actions that are observable, offering a menu of contracts is precluded. Although our APD scheme would not work perfectly in that case, there is little doubt that it would still improve the outcomes for all supply chain agents when compared with either the wholesale price contracts or with the traditional APD schemes that have been suggested in the literature.

Acknowledgments

The authors thank Serguei Netessine, Nils Rudi, George Shanthikumar, Ilia Tsetlin, Tong Wang, and seminar attendants for comments and support on various versions of this study. The authors also thank Albert Ha, the anonymous senior editor, and reviewers for their constructive and helpful feedback.

Appendix A. Proofs of the Main Results

PROOF OF LEMMA 1. Denote $D' = D|Y$. Let $F(\cdot)$ denote the cdf of the corresponding posterior distribution defined in section 4.1 and let $F_0(\cdot)$ denote the corresponding prior distribution. We adopt this notation for all the proofs unless mentioned otherwise. Without obtaining private information, for a given discount δ, if the retailer participates in the APD scheme, her profit function is $\pi = pD - (1-\delta)wq - w(D-q)^+$; otherwise, it becomes $\pi_0 = (p-w)D$. In the first case, the retailer's objective function has a typical newsvendor solution: $q_{R0} = \mu_0 + \sigma_0 z_\delta$ and corresponding profit $\mathbb{E}_D\pi = (p-(1-\delta)w)\mu_0 - w\phi(z_\delta)\sigma_0$, where $z_\delta = \Phi^{-1}(\delta)$. Note that $\mathbb{E}_D\pi \geq \mathbb{E}_D\pi_0 = (p-w)\mu_0$, that is, the retailer is always better off participating in the APD scheme.[1] On the other hand, if the retailer acquires information, given the her participation, her objective function becomes $\pi = pD' - (1-\delta)wq - w(D'-q)^+ - k$, which again has a typical newsvendor solution: $q_R = \mu + \sigma z_\delta$ and corresponding expected profit $\mathbb{E}_Y\mathbb{E}_{D'}\pi = (p-(1-\delta)w)\mu_0 - w\phi(z_\delta)\sigma - k$. The difference between the two expected profits (acquiring and not acquiring information), assuming the retailer's participation, is $w\phi(z_\delta)(\sigma_0 - \sigma) - k$, increasing in δ when $z_\delta < 0$ (i.e., $\delta < \frac{1}{2}$) and decreasing in δ otherwise. The retailer acquires information if and only if $w\phi(z_\delta)(\sigma_0-\sigma) - k \geq 0$. Hence there exist two critical discounts $0 < \underline{\delta} < \frac{1}{2} < \overline{\delta} < 1$ such that the retailer only acquires information when $\underline{\delta} \leq \delta \leq \overline{\delta}$. ☐

PROOF OF LEMMA 2. Given δ and that the retailer acquires information, the wholesaler's profit has a typical newsvendor solution with optimal order quantity $q_W = \mu + \sigma z_\gamma$ and corresponding expected profit $\mathbb{E}\pi_W = ((1-\delta)w - (1-\gamma)c)\mu_0 + w\phi(z_\delta)\sigma - c\phi(z_\gamma)\sigma$ for the retailer's corresponding order quantity $q_R = \mu + \sigma z_\delta$, where $z_\gamma = \Phi^{-1}(\gamma)$ and $z_\delta = \Phi^{-1}(\delta)$. On the other hand, the wholesaler's expected profit when APD is not offered is $\mathbb{E}\pi_{W0} = (w-(1-\gamma)c)\mu_0 - c\phi(z_\gamma)\sigma_0$. The difference between the two expected profits is $\mathbb{E}\pi_W - \mathbb{E}\pi_{W0} = -\Phi(z_\delta)w\mu_0 + w\phi(z_\delta)\sigma + c\phi(z_\gamma)(\sigma_0 - \sigma)$. This concludes the proof. ☐

PROOF OF PROPOSITION 1. Define $k^* = w\phi(z_{\delta^*})(\sigma_0 - \sigma)$, where $w\phi(z_\delta)(\sigma_0-\sigma) = k$ defines a one-to-one relationship between δ and k, for $0 < \delta \leq \frac{1}{2}$ and $k \leq \min(w\phi(0)(\sigma_0-\sigma), c\phi(z_\gamma)(\sigma_0-\sigma))$. Because of this one-to-one mapping, the problem can be equivalently cast in terms of k to make finding a solution mathematically simpler. If $k^* < k$ then the retailer would participate without acquiring private information; her resulting order quantity is depicted in Lemma 1 and her resulting benefit from participating in the APD scheme becomes $\mathbb{E}_D\pi_R - (p-w)\mu_0 = w\Phi(z_{\delta(k^*)})\mu_0 - w\phi(z_{\delta(k^*)})\sigma_0$. On the other hand, if $k^* > k$, then the retailer participates in the APD scheme and acquires information; her resulting order quantity is depicted in Lemma 1 and her resulting benefit from participating in the APD scheme becomes $\mathbb{E}_Y\mathbb{E}_{D'}\pi_R - (p-w)\mu_0 = w\Phi(z_{\delta(k^*)})$

$\mu_0 - w\phi(z_{\delta(k^*)})\sigma - k$. The benefits for the wholesaler and the supply chain can be derived in the same fashion. □

Proof of Proposition 2. We seek the optimal discount when the wholesaler offers the APD scheme. Let $\bar{k} = c\phi(z_\gamma)(\sigma_0 - \sigma)$. In response to an offer of discount δ, the retailer acquires information if and only if her benefit from acquiring information is higher than that from not (i.e., iff $w\phi(z_\delta)(\sigma_0 - \sigma) > k$). The wholesaler's expected benefit becomes $\mathbb{E}_k\Delta_W(\delta) = (-\Phi(z_\delta) w\mu_0 + w\phi(z_\delta)\sigma + \bar{k})G(w\phi(z_\delta)(\sigma_0 - \sigma)) + (-\Phi(z_\delta)w\mu_0 + w\phi(z_\delta)\sigma_0) \cdot \bar{G}(w\phi(z_\delta)(\sigma_0 - \sigma))$. Define $l = w\phi(z_\delta)(\sigma_0 - \sigma) \geq 0$, which increases in δ when $\delta \leq 0.5$ and *vice versa*. Then the wholesaler's expected benefit becomes $\mathbb{E}_k\Delta_W(\delta(l)) = w\sigma_0\phi(z_\delta) - w\mu_0\Phi(z_\delta) + (\bar{k} - l) G(l) = \frac{l\sigma_0}{\sigma_0 - \sigma} - w\mu_0\Phi(\phi^{-1}(\frac{l}{w(\sigma_0 - \sigma)})) + (\bar{k} - l)\, G(l)$. Since $w\sigma_0\phi(z_\delta) - w\mu_0\Phi(z_\delta)$ decreases in δ, for any given l, the wholesaler is always better off with the smaller discount of the two. Hence from now on we consider only $\delta \leq 0.5$. Note that here $l < \bar{k}$ is a necessary but not sufficient condition to make the wholesaler have a non-negative expected benefit since $w\sigma_0\phi(z_\delta) - w\mu_0\Phi(z_\delta) \leq 0$. The derivative of $\mathbb{E}_k\Delta_W(\delta(l))$ with respect to l becomes $\frac{d\mathbb{E}_k\Delta_W(\delta(l))}{dl} = m(l) - H(l)$, where $m(l) = \frac{\mu_0/\phi^{-1}(\frac{l}{w(\sigma_0 - \sigma)}) + \sigma_0}{\sigma_0 - \sigma}$ and $H(l) = G(l) - (\bar{k} - l)g(l)$. We therefore denote the first-order condition as

$$G(k)\left(\frac{(\bar{k} - k)g(k)}{G(k)} - 1\right) = -m(k). \tag{A1}$$

We next discuss the properties of $m(l)$ and $H(l)$ separately and how they interact with each other to form the first-order condition. To ensure a positive quantity, μ_0 and σ_0 are such that $\mu_0 + z_{\delta(l)}\sigma_0 \geq 0$, and hence $m(l) = \frac{\mu_0 + z_{\delta(l)}\sigma_0}{z_{\delta(l)}(\sigma_0 - \sigma)}$ is negative for the interested range ($\delta < \frac{1}{2}$ and hence $z_\delta < 0$); furthermore, $m(l)$ decreases in l, concave for $z_{\delta(l)} < -\sqrt{3}$ and convex otherwise; in addition, $m(0) > 0$. On the other hand, $\frac{dH(l)}{dl} = g(l)(2 - (\bar{k} - l)\frac{g'(l)}{g(l)})$ and for distributions whose pdfs are log-concave, $\frac{g'(l)}{g(l)}$ is non-increasing, which makes $2 - (\bar{k} - l)\frac{g'(l)}{g(l)}$ increasing in l; hence $2 - (\bar{k} - l)\frac{g'(l)}{g(l)}$ crosses zero at most once and $H(l)$ is either uni-modal (first decreasing and then increasing in l) or always increasing in l; in particular, when l is less than the mode, $H(l)$ is convex in l; furthermore, since as $g'(l) < 0$, $\frac{dH(l)}{dl}$ is positive for sure, the mode of $g(l)$ is larger than that of $H(l)$; in addition, $H(0) < 0$. From the above properties, we see that $m(l)$ and $H(l)$ cross at most three times, two of them when $H(l)$ decreases in l and one of them when $H(l)$ increases in l. The number of times $m(l)$ and $H(l)$ cross depends on the distribution specification and other parameters, and can be either once or three times. Then let $\mathbb{E}_k\Delta_W(\delta(l))$ reaches the potential local maximum, minimum, and maximum at l_1, l_0, and l_2 respectively, where $l_1, l_0, l_2 = \{l | m(l) - H(l) = 0\}$. The maximum occurs when $m'(l) - H'(l) < 0$; in particular, when $H'(l) > 0$ or when $H'(l) < 0$ but the slope of $m(l)$ is steeper than that of $H(l)$, hence the three solutions to the first-order condition should be in the order of a maximum, minimum, and a maximum as l increases; consequently, let $l_1 < l_0 < l_2$. Therefore, when there is only one solution to the first-order condition $m(l) - H(l) = 0$, it is either l_1 or l_2 and k^* should equal that value; when there are three solutions to the first-order condition, we would take the smallest and the largest of these three solutions, that is, we take l_1 and l_2 and then $k^* = \text{argmax}\{\mathbb{E}_k\Delta_W(\delta(l_1)), \mathbb{E}_k\Delta_W(\delta(l_2))\}$.

Denote by k_{0i}, $i = 1, 2$, the solution of Equation (A1) for the distribution with pdf $g_i(k)$ and cdf $G_i(k)$. Define l_1 and l_2 as any solution to the first-order condition $m(l) - H_1(l) = 0$ and $m(l) - H_2(l) = 0$ respectively, where $H_i(l) = G_i(l) - (\bar{k} - l)g_i(l)$, $i = 1, 2$. $g_1(l)(\bar{k} - l) - G_1(l) < g_2(l)(\bar{k} - l) - G_2(l)$ corresponds to $H_1(l) > H_2(l)$, hence we have

$$0 = m(l_1) - H_1(l_1) < m(l_1) - H_2(l_1) > m(l_2) - H_2(l_2) = 0,$$

where the two equalities are by definition of l_1 and l_2. The first inequality comes from the assumption that for a given l, $H_1(l) < H_2(l)$, and the second has to hold since otherwise we would have $0 < 0$. From the second inequality, since $m(l) - H(l)$ decreases in l (due to concavity), we therefore see that $l_2 > l_1$ and consequently, the corresponding $k_2^* \geq k_1^*$ and $\delta(k_2^*) \geq \delta(k_1^*)$. □

Proof of Proposition 3. Observe that $g(k|u,s) = \frac{1}{s}g_0\left(\frac{k-u}{s}\right)$. From Equation (A1), we have

$$m(k^*) - G_0\left(\frac{k^* - u}{s}\right) + \frac{1}{s}(\bar{k} - k^*)g_0\left(\frac{k^* - u}{s}\right) = 0. \tag{A2}$$

Define $s_0 = \frac{(\bar{k} - u)g_0(0)}{G_0(0) - m(u)}$. We will show that $k^* \geq u$ when $s \leq s_0$ and otherwise when $s > s_0$: (i) when $s \leq s_0$, we assume the opposite, letting $k^* < u$, then $\frac{k^* - u}{s} < 0$ and given that $m(l) - H(l)$ decreases in l, from Equation (A2) we see that $0 = m(k^*) + \frac{1}{s}(\bar{k} - k^*)g_0(\frac{k^* - u}{s}) - G_0(\frac{k^* - u}{s}) > m(u) + \frac{1}{s}(\bar{k} - u) g_0(0) - G_0(0) > m(u) + \frac{1}{s_0}(\bar{k} - u)g_0(0) - G_0(0) = 0,$

a contradiction; (ii) when $s > s_0$, we let $k^* > u$ instead, then $\frac{k^*-u}{s} > 0$ and given that $m(l) - H(l)$ decreases in l, from Equation (A2) we see that $0 = m(k^*) + \frac{1}{s}(\bar{k} - k^*)g_0(\frac{k_0-u}{s}) - G_0(\frac{k^*-u}{s}) < m(u) + \frac{1}{s}(\bar{k} - u)g_0(0) - G_0(0) < m(u) + \frac{1}{s_0}(\bar{k} - u)g_0(0) - G_0(0) = 0$, another contradiction. Then we prove (i). We look at the first-order condition where $-H(l)$ decreases in l. First of all, we show that when $s \geq s_0$, k^* decreases in s. Since $k^* < u$ when $s \geq s_0$, $\frac{k^*-u}{s}$ increases in s, which makes $\frac{1}{s}(\bar{k} - k^*)g_0(\frac{k^*-u}{s}) - G_0(\frac{k^*-u}{s})$ decreasing in s. To balance the equality of Equation (A2), k^* has to decrease. Hence, when $s \geq s_0$, k^* decreases in s. Next, we observe that when $s = 0$ or when $s = s_0$, $k^* = u$. Hence when $s < s_0$, k^* should be first increasing and then decreasing in s at least once. We define $k_\downarrow$ and $k_\uparrow$ the two solutions to the first-order condition where $H(l)$ decreases and increases in l respectively; in particular $k_\downarrow < u$. When $s = 0$, the wholesaler knows the retailer's cost for sure and hence $k^* = k_\uparrow = u$, that is, the expected benefit for the wholesaler is higher at $k_\uparrow$ than at $k_\downarrow$. We write the difference of the two expected benefits as $\mathbb{E}_k\Delta_W(\delta(k_\uparrow)) - \mathbb{E}_k\Delta_W(\delta(k_\downarrow)) = n(k_\uparrow) - n(k_\downarrow) + (\bar{k} - k_\uparrow)\,G(k_\uparrow) - (\bar{k} - k_\downarrow)G(k_\downarrow) = n(k_\uparrow) - n(k_\downarrow) + (\bar{k} - k_\uparrow) G(\frac{k_\uparrow-u}{s}) - (\bar{k} - k_\downarrow)G(\frac{k_\downarrow-u}{s})$, where $n(l) = w\sigma_0\phi(z_{\delta(l)}) - w\mu_0\Phi(z_{\delta(l)})$; when $s = 0$, such difference is positive. By (ii), $k_\uparrow$ first increases in s from $k_\uparrow = u$, hence $k_\uparrow \geq u$ whereas $k_\downarrow < u$; as s increases, $\frac{k_\uparrow-u}{s}$ decreases whereas $\frac{k_\downarrow-u}{s}$ increases, which shortens the difference between the two expected benefits and eventually might switch the sign such that $k^* = k_\downarrow < u$. □

Proof of Corollary 1. The corollary follows from the benefits listed in Proposition 1—namely, $k^*(u, s) \geq u$ when $s \leq s_0$ and $k^*(u, s) < u$ when $s > s_0$. □

Note

[1]Note that under normal distribution this holds when the demand is non-negative, which can be more or less ensured when the coefficient of variation $\frac{\sigma_0}{\mu_0}$ is relatively small, e.g., $\frac{\sigma_0}{\mu_0} < \frac{1}{3}$ (Gallego 1995, Li and Zhang 2008). Importantly, this can be shown for a general demand distribution. We have that $\mathbb{E}_D\pi = p\mu_0 - (1 - \delta)wq - w\int_q^\infty (D - q)dF$ and $\mathbb{E}_D\pi_0 = (p - w)\mu_0$. The newsvendor solution is now derived from $\bar{F}(q_{R0}) = 1 - \delta$. Then the difference between the two profits becomes $\mathbb{E}_D\pi|_{q_{R0}} - \mathbb{E}_D\pi_0 = -(1 - \delta)wq + w\int_0^q DdF + wq\bar{F}(q)\big|_{q_{R0}} = w\int_0^{q_{R0}} DdF \geq 0$. Note that this inequality is binding when $q_{R0} = 0$.

References

Aiyer, S., G. Ledesma. 2004. Waste not, want not. *Log. Today* **45**(4): 35–37.

Boyaci, T., Ö. Özer. 2010. Information acquisition for capacity planning via pricing and advance selling: When to stop and act? *Oper. Res.* **58**(5): 1328–1349.

Cachon, G. 2004. The allocation of inventory risk in a supply chain: Push, pull, and advance-purchase discount contracts. *Management Sci.* **50**(2): 222–238.

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

Cachon, G., M. A. Lariviere. 2005. Supply chain coordination with revenue-sharing contracts: strengths and limitations. *Management Sci.* **51**(1): 30–44.

Cachon, G., C. Terwiesch. 2006. *Matching Supply with Demand: An Introduction to Operations Management*. McGraw-Hill, Primis Custom Publishing, New York City, NY.

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

Cho, S. H., C. S. Tang. 2013. Advance selling in a supply chain under uncertain supply and demand. *Manuf. Serv. Oper. Manag.* **15**(2): 305–319.

Corbett, C. J. 2001. Stochastic inventory systems in a supply chain with asymmetric information: cycle stocks, safety stocks, and consignment stock. *Oper. Res.* **49**(4): 487–500.

Corbett, C. J., X. de Groote. 2000. A supplier's optimal quantity discount policy under asymmetric information. *Management Sci.* **46**(3): 444–450.

Corbett, C. J., D. Zhou, C. S. Tang. 2004. Designing supply contracts: Contract type and information asymmetry. *Management Sci.* **50**(4): 550–559.

Dai, T., S. H. Cho, F. Zhang. 2016. Contracting for on-time delivery in the U.S. influenza vaccine supply chain. *Manuf. Serv. Oper. Manag.* **18**(3): 332–346.

Daughety, A. F., J. F. Reinganum. 1994. Asymmetric information acquisition and behavior in role choice models: An endogenously generated signaling game. *Int. Econ. Rev.* **35**(4): 795–819.

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

Fisher, M., A. Raman. 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* **44**(1): 87–99.

Fisher, M., J. H. Hammond, W. R. Obermeyer, A. Raman. 1994. Making supply meet demand in an uncertain world. *Harvard Bus. Rev.* **72**(3): 83–93.

Gabaix, X., D. Laibson, G. Moloche, S. Weinberg. 2006. Costly information acquisition: Experimental analysis of a boundedly rational model. *Am. Econ. Rev.* **96**(4): 1043–1068.

Gallego, G. 1995. The newsvendor problem. Lecture notes, Columbia University, New York, NY.

Girotra, K., W. Tang. 2009. Strategic behavior in supply chains: Information acquisition. C. Tang, S. Netessine, eds. *Consumer-Driven Demand and Operations Management Models*, vol. 131, chap. 18. Springer, Berlin, Germany, 459–477.

Girotra, K., W. Tang. 2010. Costume Gallery Inc. INSEAD Teaching Case.

Gundepudi, P., N. Rudi, A. Seidmann. 2001. Forward vs. spot buying of information goods onWeb: Analyzing the consumer decision process. *J. Manage. Inf. Syst.* **18**(2): 107–131.

Guo, L. 2009. The benefits of downstream information acquisition. *Management Sci.* **28**(3): 457–471.

Gurnani, H., C. S. Tang. 1999. Optimal ordering decisions with uncertain cost and demand forecast updating. *Management Sci.* **45**(10): 1456–1462.

Ha, A. 2001. Supplier-buyer contracting: Asymmetric cost information and cut-off level policy for buyer participation. *Nav. Res. Log.* **48**(1): 41–64.

Hammond, J. H., A. Raman. 1994. Sport Obermeyer, Ltd. Harvard Business School Case.

Hays, C. L. 2004. What Wal-Mart knows about customers habits. Available at http://www.nytimes.com/2004/11/14/business/yourmoney/what-walmart-knows-about-customers-habits.html?_r=0 (accessed date September 30, 2015).

Health Sciences Authority. 2009. Understanding vaccines, vaccine development and production. Technical report, Health Sciences Authority of Singapore Government.

Heinrich, J. 2001. Flu vaccine: Supply problems heighten need to ensure access for high-risk people. Technical report, GAO-01-624, 15 May, Washington, DC.

Kalkanci, B., K. Y. Chen, F. Erhun. 2011. Contract complexity and performance under asymmetric demand information: An experimental evaluation. *Management Sci.* **57**(4): 689–704.

Katok, E., V. Pavlov. 2013. Fairness in supply chain contracts: A laboratory study. *J. Oper. Manag.* **31**(3): 129–137.

Katok, E., T. Olsen, V. Pavlov. 2014. Wholesale pricing under mild and privately known concerns for fairness. *Prod. Oper. Manag.* **23**(2): 285–302.

Kurtulus, M., S. Ülkü, B. L. Toktay. 2012. The value of collaborative forecasting in supply chains. *Manuf. Serv. Oper. Manag.* **14**(1): 82–98.

Li, L., H. Zhang. 2008. Confidentiality and information sharing in supply chain coordination. *Management Sci.* **54**(8): 1467–1481.

Li, C., F. Zhang. 2013. Advance demand information, price discrimination, and preorder strategies. *Manuf. Serv. Oper. Manag.* **15**(1): 57–71.

Lovejoy, W. S. 2010. Conversations with supply chain managers. Working paper, Ross Business School, University of Michigan.

McCardle, K., K. Rajaram, C. S. Tang. 2004. Advance booking discount programs under retail competition. *Management Sci.* **50**(5): 701–708.

O'Mara, D., K. Fukuda, J. A. Singleton. 2003. Influenza vaccine: Ensuring timely and adequate supplies. *Infect. Med.* **20**(11): 548–554.

Özer, Ö., W. Wei 2006. Strategic commitments for an optimal capacity decision under asymmetric forecast information. *Management Sci.* **52**(8): 1238–1257.

Özer Ö., O. Uncu, W. Wei. 2007. Selling to the newsvendor with a forecast update. *Eur. J. Oper. Res.* **182**(3): 1150–1176.

Prasad, A., K. E. Stecke, X. Zhao. 2011. Advance selling by a newsvendor retailer. *Prod. Oper. Manag.* **20**(1): 129–142.

Rosling, K. 2002. Inventory cost rate functions with nonlinear shortage costs. *Oper. Res.* **50**(6): 1007–1017.

Rudi, N., D. Drake. 2014. Observation bias: The impact of demand censoring on newsvendor level and adjustment behavior. *Management Sci.* **60**(5): 1334–1345.

Shugan, S. M., J. Xie. 2004. Advance selling for services. *Calif. Manage. Rev.* **46**(3): 37–54.

Tang, C. S., K. Rajaram, A. Alptekinoglu, J. Ou. 2004. The benefits of advance booking discount programs: Model and analysis. *Management Sci.* **50**(4): 465–478.

Taylor, T. A., W. Xiao. 2009. Incentives for retailer forecasting: Rebates versus returns. *Management Sci.* **55**(10): 1654–1669.

Williams, D. G. 2005. The influenza vaccine supply chain: Structure, risk and coordination. Masters thesis. Massachusetts Institute of Technology, Cambridge, MA.

Winkler, R. L. 1981. Combining probability distributions from dependent information sources. *Management Sci.* **27**(4): 479–488.