Does a Manufacturer Benefit from Selling to a Better-Forecasting Retailer?

Terry A. Taylor and Wenqiang Xiao

Haas School of Business, University of California, Berkeley

Stern School of Business, New York University

Abstract

This paper considers a manufacturer selling to a newsvendor retailer that possesses superior demand-forecast information. We show that the manufacturer's expected profit is convex in the retailer's forecasting accuracy: The manufacturer benefits from selling to a better-forecasting retailer if and only if the retailer is already a good forecaster. If the retailer has poor forecasting capabilities, then the manufacturer is *hurt* as the retailer's forecasting capability improves. More generally, the manufacturer tends to be hurt (benefit) by improved retailer forecasting capabilities if the product economics are lucrative (poor). These results hold under both the optimal procurement contract and the optimal wholesale price contract. Further, the optimal procurement contract is a quantity discount contract.

Key words: supply-chain contracting; asymmetric information; forecasting

November 2008

1 Introduction

Because they are closer to the end customer, retailers often have better information about market demand than the manufacturer. However, the degree of superiority in forecasting demand varies (1) across retailers and (2) over time. First, some retailers are known to be better forecasters than others. For example, in the retail consumer electronics industry, Circuit City has been plagued by a weak forecasting capabilities and trails behind best-in-class retailer Best Buy (Feldman and Cramer 2004, Widlitz 2005). How does a retailer's effectiveness in forecasting influence her attractiveness to a manufacturer that is selecting a retail partner? When faced with a pool of prospective retailers, *ceteris paribus,* should a manufacturer select a retailer that has strong, weak, or intermediate forecasting capabilities?

Not only do forecasting capabilities vary across retailers, they also vary over time at a single retailer. For example, a manufacturer's retail partner may invest in forecasting capabilities (e.g., by purchasing relevant software) with the intention of improving its forecasting accuracy. Alternately, the retailer may disinvest in forecasting (e.g., by laying off or redeploying forecasting staff), understanding that this action will degrade its ability to accurately forecast demand. What impact should the manufacturer anticipate that such changes by its retail partner will have on the manufacturer's own performance? Should a manufacturer relish and encourage either improved or worsened retailer forecasting accuracy?

It may seem natural that a manufacturer would benefit by selling to a better-forecasting retailer in that by doing so, production decisions can be made with more accurate information, reducing the cost of supply/demand mismatch. However, selling to a better-informed retailer puts the manufacturer at a strategic disadvantage relative to the retailer. The retailer may be able to use her informational advantage to extract a larger portion of the system profit from the manufacturer. The impact of improved retailer forecasting accuracy on manufacturer profit depends on the trade-off between these two factors, and it is this trade-off that we explore in this paper.

This paper considers a manufacturer selling to a newsvendor retailer that possesses superior demand-forecast information. We show that the manufacturer's expected profit is convex in the retailer's forecasting accuracy: The manufacturer benefits from selling to a better-forecasting retailer if and only if the retailer is already a good forecaster. If the retailer has poor forecasting capabilities, then the manufacturer is *hurt* as the retailer's forecasting capability improves. More generally, the manufacturer tends to be hurt (benefit) by improved retailer forecasting capabilities if the product economics are lucrative (poor). These results hold under both the optimal procurement contract and the optimal wholesale price contract. Further, the optimal procurement contract is a quantity discount contract.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature. Section 3 describes the model. Sections 4, 5, and 6 contain the analysis for the integrated solution, the wholesale price contract, and the general procurement contract, respectively. Section 7 provides numerical results. Section 8 provides evidence for the robustness of our main result. Section 9 provides concluding remarks. All proofs are in the appendix, with the exception that the proof of the last result is in Taylor and Xiao (2008).

2 Literature Review

There is substantial literature that studies supply chain settings in which firms have distinct demand information. This literature can be classified into two streams. One stream considers the impact of the truthful sharing of private demand information (e.g., Cachon and Fisher 2000, Lee et al. 2000, and Aviv 2001). Demand information sharing sometimes is not difficult to achieve, e.g., when demand information only consists of historical sales data that are readily verifiable from the information system. In this case, the interesting questions include how to use the shared information to improve supply chain performance and what factors are crucial in affecting the magnitude of the improved performance. However, when demand information also involves the firms' subjective assessment or private knowledge that is not verifiable by a third party, the credibility of truthful information sharing is in doubt because a firm may have incentive to misrepresent its information. (For example, in our setting, the privately-informed retailer has incentive to persuade the manufacturer that market demand is weak (regardless of the actual market condition), so as to convince the manufacturer to offer more attractive terms (e.g., a lower wholesale price).) Indeed, in practice, the scope for opportunistic behavior and lack of trust have proven to be substantial obstacles to demand forecasting collaboration efforts (Fliedner 2003).

To counter this credibility issue, a second stream of literature studies how to design incentive contracts so that it is in the firm's best interest to truthfully share its demand information (Cachon and Lariviere 2001, Arya and Mittendorf 2004, Özer and Wei 2006, Burnetas et al. 2007, and Ren et al. 2008). A typical theme in this stream of work is to explore how contracts should be designed and then to evaluate the performance of optimal contracts and/or simple and commonly-used contracts. See Cachon (2003) and Chen (2003) for reviews. Our work fits within this stream. However, the focus of our work is distinct: We concentrate on the impact of the retailer's forecasting accuracy on the firms' performance.

The paper most closely related to our work is Miyaoka and Hausman (2008). Similar to our work, they study a supply chain where the upstream firm (supplier) sells to a downstream newsvendor (manufacturer) who has private demand-forecast information. Unlike ours, the upstream firm must make a capacity decision. Even so, we share the same objective, which is to evaluate the impact of the downstream firm's forecasting accuracy on the firms' performance. However, they restrict attention to the single wholesale price contract, while we study the issue under both the wholesale price contract and the optimal procurement contract, with the emphasis on the latter. Further, our results for the wholesale price contract generalize theirs in the sense that they obtain analytical results for the two extreme cases of forecasting accuracy (the downstream firm is either perfectly informed or completely uninformed) and provide numerical results for intermediate cases, while we obtain analytical results for the full spectrum. In particular, we provide a more complete characterization as to how the upstream firm's performance changes in the downstream firm's forecasting accuracy (e.g., the convexity property). This issue of studying the impact of the forecasting accuracy on supply chain performance has also been discussed, mainly through numerics, in Özer and Wei (2006) and Taylor (2006).

Our work is also related to the economics/accounting literature that studies the optimal level of information asymmetry. Lewis and Sappington (1991) and Rajan and Saouma (2006) consider a principal-agent model where the agent privately exerts effort that influences the output. The agent has private information about his cost of effort. The authors examine the impact of the accuracy of the agent's private information on the principal's utility. They establish an "all-or-nothing" result: The principal prefers to deal with either a completely-uninformed agent or a perfectly-informed agent. Even though our supply chain setting with asymmetric demand information is quite different, it is interesting that we also obtain this "all-or-nothing" result.

3 Model

A manufacturer (he) produces a product at unit cost c and sells to a newsvendor retailer (she), who then sells at a fixed retail price p to a market with random demand D in a single selling season. The market demand is normally distributed, i.e., $D \sim N(\mu_0, \sigma_0)$. The salvage value of unsold inventory is normalized to zero.

Prior to the selling season, the retailer observes a demand forecast $S = D + \varepsilon$, where $\varepsilon \sim N(0, \sigma_1)$ is independent of D. Note that S is an unbiased estimator of D, with the estimation error being normally distributed. It follows from the conjugate property of normal distribution that the posterior demand distribution under the forecast S is also normal (see Winkler 1981), i.e.,

$$D|_S \sim N(a^2\mu_0 + (1 - a^2)S, a\sigma_0), \tag{1}$$

where

$$a \equiv \frac{\sigma_1}{\sqrt{\sigma_0^2 + \sigma_1^2}}$$

denotes the fraction of the original demand uncertainty, as measured by the standard deviation, that remains after the forecast is observed. We refer to a as the retailer's *forecasting accuracy.* The lower the value of a, the more accurate the retailer's forecast. In the limiting case where $a = 0$, the forecast perfectly reveals the exact demand. In the opposite limiting case where $a = 1$, the forecast contains no valuable information about demand and the posterior distribution is identical to the prior. For expositional simplicity, we exclude these two extreme cases and restrict attention to $a \in (0, 1)$. In considering the impact of changes in forecasting accuracy, we assume that the distribution of the underlying demand D (i.e., the parameters μ_0 and σ_0) is fixed and only the level of noise in the retailer's forecast (as captured in σ_1) varies.

The retailer privately observes her own forecast. However, the distributions and all other parameters are common knowledge of both the manufacturer and the retailer.[1] Thus the manufacturer

---

[1] Typically, a manufacturer has some awareness of its (current or prospective) retailer's demand-forecasting investments in computer systems, software, and staff. Further, the manufacturer may have an understanding of the retailer's historic forecasting performance (either directly, or reflected in the retailer's history of stocking too much or too little) and the retailer's familiarity with the manufacturer's product (based, for example, on what products the retailer has

knows that the retailer's forecast is normally distributed, i.e., $S \sim N(\mu_0, \sqrt{\sigma_0^2 + \sigma_1^2})$, or equivalently, $S \sim N(\mu_0, \sigma_0/\sqrt{1-a^2})$. It is convenient to rewrite $S = \mu_0 + \left(\sigma_0/\sqrt{1-a^2}\right)\Theta$, where $\Theta \sim N(0,1)$. Because there is a one-to-one mapping between S and Θ, we also refer to Θ as the retailer's forecast. Given the retailer's forecast $\Theta = \theta$, the posterior distribution in (1) is

$$D|_{\Theta=\theta} \sim N(\mu_0 + \sqrt{1-a^2}\sigma_0\theta, a\sigma_0). \tag{2}$$

Note that the retailer's forecast accuracy a impacts both the mean and standard deviation of posterior demand given a forecast: The more accurate the forecast (smaller a), the larger the weight of forecast ($\sqrt{1-a^2}\sigma_0$) in determining the posterior mean, and the smaller the posterior standard deviation ($a\sigma_0$).

We assume both the manufacturer and the retailer are risk neutral, maximizing their own expected profits. Because the retailer possesses superior information about the demand, the uninformed manufacturer faces a typical adverse selection problem in contracting with the informed retailer. In such a situation, a procurement contract can be represented by a transfer payment schedule $T(q, \theta)$, which specifies the payment the retailer makes to the manufacturer when the retailer orders q units and reports observing forecast θ. The commonly-used wholesale price contract is a special case of this general form: $T(q, \theta) = wq$, where the contract specification reduces to a single parameter, the wholesale price w.

The sequence of events is as follows: First, the manufacturer specifies a transfer payment schedule $T(q, \theta)$ (without knowing the retailer's forecast), and the retailer (privately) observes the forecast Θ. Second, the retailer orders q units and reports a forecast $\hat{\theta}$; the manufacturer fulfills the retailer's order and receives the payment $T(q, \hat{\theta})$ from the retailer. Third, the market demand D is realized, and the retailer receives sales revenue $p \min(D, q)$.

Our goal is to examine the impact of the retailer's forecasting accuracy on the manufacturer's performance under the optimal procurement contract and under the optimal wholesale price contract.

---

carried in the past). To the extent that this awareness and understanding is reasonably good, the assumption that the manufacturer knows or can infer the retailer's forecasting accuracy is reasonable, at least as an approximation. However, in some cases the manufacturer may lack an understanding of the retailer's forecasting capabilities (e.g., if the manufacturer is unfamiliar with the retailer and relevant public information is scarce); to address this case, a more complex model capturing this additional dimension of private information is required.

4 Integrated System

As a benchmark, we first briefly examine the impact of the retailer's forecasting accuracy on the performance of integrated system, where there is a single decision maker. After observing the demand forecast $\Theta = \theta$, the system faces the demand $D|_{\Theta=\theta}$ with distribution (2). Letting $D_\theta \equiv D|_{\Theta=\theta}$, we can write

$$D_\theta = \mu_0 + \sqrt{1-a^2}\sigma_0\theta + a\sigma_0 X, \tag{3}$$

where $X \sim N(0,1)$. Let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal density and distribution function, respectively, and $\overline{\Phi}(\cdot) \equiv 1 - \Phi(\cdot)$. The system decides how much to produce by solving the following newsvendor problem, denoted by (**P1**),

$$(\mathbf{P1}) \max_q \; E_{D_\theta}[p\min(D_\theta, q) - cq],$$

which can be rewritten as (see, e.g., Porteus 2002)

$$\max_z \; \{(p-c)(\mu_0 + \sqrt{1-a^2}\sigma_0\theta) - a\sigma_0\Gamma(z)\}, \tag{4}$$

where $z \equiv (q - \mu_0 - \sqrt{1-a^2}\sigma_0\theta)/(a\sigma_0)$ and $\Gamma(z) \equiv p[\phi(z) - z\overline{\Phi}(z)] + cz$. The first term in (4) is the profit on the mean demand, and the second term is the expected cost of supply/demand mismatch with order quantity $q = \mu_0 + \sqrt{1-a^2}\sigma_0\theta + a\sigma_0 z$:

$$a\sigma_0\Gamma(z) = (p-c)E_{D_\theta}[D_\theta - q]^+ + cE_{D_\theta}[q - D_\theta]^+,$$

where $x^+ \equiv \max(x, 0)$. Because $\Gamma(z)$ is strictly convex and is minimized at $z^I = \overline{\Phi}^{-1}(c/p)$, given the forecast $\Theta = \theta$, the system's optimal production quantity is

$$q^I(\theta) = \mu_0 + \sqrt{1-a^2}\sigma_0\theta + a\sigma_0 z^I, \tag{5}$$

and the system's expected profit is

$$\Pi^I(\theta) = (p-c)(\mu_0 + \sqrt{1-a^2}\sigma_0\theta) - a\sigma_0\Gamma(z^I). \tag{6}$$

Because Θ is standard normal, the system's expected profit is

$$\Pi^I = E_\Theta \Pi^I(\Theta) = (p-c)\mu_0 - a\sigma_0\Gamma(z^I).$$

Intuitively, the integrated system benefits from a higher forecasting accuracy (i.e., smaller a): With better forecast information, the system makes a better-informed production quantity decision, which reduces the cost of supply-demand mismatch. In our case this is manifest by the fact that Π^I decreases in a, a consequence of the well-known result that under normal demand, a newsvendor's expected profit decreases in the standard deviation of demand. Further, $(d/da)\Pi^I = -\sigma_0\Gamma(z^I) = -\sigma_0 p\phi(z^I)$, and $p\phi(z^I)$ is increasing in p and increasing in c on $c \in (0, p/2)$ and decreasing in c on $c \in (p/2, p)$ (Qi and Zhu 2006). The implication is that improvement in forecasting accuracy is of greater value to the integrated system when the price p is high, the cost c is moderate (close to $p/2$), and the underlying demand is volatile (σ_0 is large). Because the impact of forecasting accuracy on an integrated system is well-established, our contribution is in the analysis of the decentralized system, to which we turn in the next two sections.[2]

5 Wholesale Price Contract

Although the integrated system always benefits from improved forecasting accuracy, it is not clear whether, in the decentralized system, the manufacturer will always benefit from selling to a retailer with a higher forecasting accuracy. Although the focus of this paper is the general procurement contract $T(q, \theta)$, to build intuition and for completeness, we begin by considering the simple wholesale price contract $T(q, \theta) = wq$, whereby the manufacturer sells to the retailer at wholesale price w.

Given w and forecast $\Theta = \theta$, the retailer decides how much to order to maximize her expected profit. The retailer's problem is identical to the integrated system's problem (**P1**) with c being replaced by w. Therefore, as the counterparts of (5) and (6), the retailer's optimal order quantity is

$$q(w, \theta) = \mu_0 + \sqrt{1-a^2}\sigma_0\theta + a\sigma_0 z, \tag{7}$$

where $z = \overline{\Phi}^{-1}(w/p)$, and her expected profit is

$$R(w, \theta) = (p - w)(\mu_0 + \sqrt{1-a^2}\sigma_0\theta) - a\sigma_0[\Gamma(z) + (w - c)z]. \tag{8}$$

Now we turn to the manufacturer's problem. Without knowing the realized forecast value, the manufacturer sets the wholesale price w to maximize his expected profit, where the expectation

---

[2] In the decentralized system, inefficiency arises because, although the manufacturer offers the contract, it is the retailer's right to choose the order quantity. If this decision right of the retailer can be transferred to the manufacturer, then the manufacturer is essentially transformed into a centralized decision-maker; the manufacturer achieves the integrated system profit by asking the retailer for her forecast and then dictating that the retailer order the integrated system optimal quantity; and the impact of forecasting accuracy on the manufacturer's expected profit is identical to its impact on the integrated system's.

is taken over the retailer's forecast Θ. Consequently, the manufacturer's contract design problem, denoted by (**P2**), is

$$(\mathbf{P2})\ \max_{w \geq c}(w-c)E_{\Theta}q(w,\Theta).$$

Because $w = p\overline{\Phi}(z)$, there is a one-to-one mapping between w and z. Therefore, we can rewrite (**P2**) as

$$\max_{z \leq \overline{\Phi}^{-1}(c/p)} \quad (p\overline{\Phi}(z)-c)(\mu_0 + a\sigma_0 z).$$

Let z^* denote the solution to this optimization problem; z^* is the number of standard deviations of safety stock purchased by the retailer under the optimal contract. The manufacturer's optimal wholesale price is $w^* = p\overline{\Phi}(z^*)$, and the manufacturer's expected profit is $M^w = (p\overline{\Phi}(z^*) - c)(\mu_0 + a\sigma_0 z^*)$.

Next we examine how the retailer's forecasting accuracy a impacts the manufacturer's expected profit M^w. To build intuition, we first describe how a impacts z^* and w^*.

Lemma 1. *z^* (w^*) is strictly increasing (decreasing) in the retailer's forecasting accuracy a (Lariviere and Porteus 2001).*

As the retailer's forecast accuracy improves (a decreases), it is optimal for the manufacturer to charge a higher wholesale price. This result is intuitive because when the retailer is very confident about her forecast, she will order a quantity that is close to her forecast regardless of the wholesale price; consequently, the manufacturer can get away with charging a high wholesale price. Let $a^w \equiv p\mu_0/[\sqrt{2\pi}(p/2-c)\sigma_0]$. We now turn to the main result of this section.

Proposition 1. *The manufacturer's profit under the optimal wholesale price contract, M^w, is strictly convex in the retailer's forecasting accuracy a. If $c < p/2$ and $a^w < 1$, M^w strictly decreases in a for $a \in (0, a^w)$ and strictly increases for $a \in (a^w, 1)$. Otherwise, M^w decreases in a for $a \in (0,1)$.*

Whether the manufacturer benefits from improved retailer forecasting accuracy depends on the retailer's current forecasting accuracy a. In general, there are two regimes, $a \in (0, a^w)$ and $a \in (a^w, 1)$, although, as discussed below, these collapse into a single regime for a class of model parameters. The manufacturer benefits from the improved retailer forecasting accuracy if $a \in (0, a^w)$, i.e., the retailer's current forecasting accuracy is already very good; while the opposite is true if $a \in (a^w, 1)$, i.e., the retailer's current forecasting accuracy is very poor.

The intuition for the first result is as follows. When the retailer is good at forecasting, she is relatively certain about the actual demand based on her forecast, and so her purchase quantity is relatively insensitive to the wholesale price (see equation (7)). This makes it attractive for the manufacturer to charge the retailer a high wholesale price ($w^* > p/2$). Under the high wholesale price, the retailer is conservative in choosing her order quantity: She orders less than the posterior mean demand, i.e., $z^* < 0$. Improved retailer forecasting accuracy pushes the retailer to order closer to the posterior mean (i.e., less conservatively), which benefits the manufacturer.

The intuition for the second result is as follows. When the retailer's forecasting accuracy is poor, she is uncertain about the demand, and so her purchase quantity is sensitive to the wholesale price. This makes it attractive for the manufacturer to offer a low wholesale price ($w^* < p/2$) because the quantity-effect more than makes up for the reduced margin. Under the low wholesale price, the retailer is aggressive in choosing her order quantity: She orders more than the posterior mean demand, i.e., $z^* > 0$. Improved retailer forecasting accuracy pushes the retailer to order closer to the posterior mean (i.e., less aggressively), which hurts the manufacturer.

The regime in which this second result occurs, $a \in (a^w, 1)$, exists only when the production cost is small and the uncertainty in the underlying demand is large ($c \ll p$ and high coefficient of variation σ_0/μ_0 imply $a^w < 1$). When the production cost is large or the uncertainty in the underlying demand is small, it is attractive for the manufacturer to charge a high wholesale price ($w^* > p/2$), and so, for the reasons described above, the manufacturer benefits by improved retailer forecasting accuracy.

To explore the implications of the convexity result in Proposition 1, consider a manufacturer that is selecting a retailer to distribute his product from a pool of N retailers. Retailer i has forecasting accuracy a_i, and $0 < a_1 < a_2 < ... < a_N < 1$. Because the manufacturer's expected profit is convex in a, it is optimal for the manufacturer to select the best ($a = a_1$) or worst ($a = a_N$) forecaster. It is easy to check that as $a \to 0$, $M^w \to (p - c)\mu_0$, which is equal to the integrated system expected profit Π^I with $a = 0$. This is clearly the upper bound on the maximum expected profit that the manufacturer could possibly achieve. Therefore, if a_1 is sufficiently close to 0, then the manufacturer's optimal choice is to contract with the best forecaster ($a = a_1$). However, it is possible that the manufacturer may prefer to contract with the worst forecaster ($a = a_N$) when the best-forecasting retailer is fairly weak (e.g., $a_1 \geq a^w$).

The implication is that manufacturers ought to avoid blindly seeking out retailers with strong forecasting capabilities. If the production cost is high or demand uncertainty is low, then this naively-appealing approach will serve the manufacturer well. However, if the production cost is low and demand uncertainty is high, the manufacturer may benefit by selling to a retailer with inferior forecasting capabilities. We are not the first to point out that the manufacturer may benefit by selling to a retailer with inferior forecasting capabilities. Taylor (2006) and Miyaoka and Hausman (2008) provide numerical examples in which the manufacturer's expected profit is decreasing and then increasing in the retailer's forecasting accuracy; we complement this work by establishing the convexity result analytically.

To summarize, this section's main qualitative insight is that the manufacturer's expected profit first decreases and then increases as the retailer's forecasting accuracy improves. This result was established under the optimal wholesale price contract, and one might conjecture that the result is an artifact of that particular and simplistic contractual form. In the next section we show that, to the contrary, this main qualitative result continues to hold under the optimal procurement contract.

6 General Procurement Contract

In the previous section, the transfer payment was restricted to be a linear function of the retailer's order quantity. However, in practice, it is not uncommon for procurement contracts to specify that the unit price depends on the quantity purchased (i.e., the transfer payment is nonlinear). In this section we consider general procurement contracts, in which the transfer payment is allowed to be of a general (i.e., nonlinear) form. First, we characterize the optimal procurement contract (Propositions 2 and 3). Second, we characterize the impact of the retailer's forecasting accuracy on the manufacturer's profit under this optimal contract (Proposition 4).

In our model setting with asymmetric information, a general procurement contract can be represented by a transfer payment $T(q, \theta)$, which is a function of the retailer's order quantity q and reported forecast θ. It follows from the revelation principle that in our model setting, finding the optimal transfer payment function is equivalent to finding the optimal menu of quantity-payment pairs $\{q(\theta), t(\theta)\}$ that induces retailer truth-telling. The retailer selects from the menu by "reporting" a forecast $\hat{\theta}$, which corresponds to selecting the contract that stipulates $q(\hat{\theta})$ units as the purchase quantity and $t(\hat{\theta})$ as the transfer payment. The menu induces truth-telling if it is in the retailer's interest to report the forecast she actually observed.

The retailer observing a forecast $\Theta = \theta$ is referred to as the type-θ retailer. This retailer faces demand D_θ. If the type-θ retailer chooses the quantity-payment pair $(q(\hat{\theta}), t(\hat{\theta}))$, then her expected profit is

$$R(\theta, \hat{\theta}) = pE_{D_\theta} \min(D_\theta, q(\hat{\theta})) - t(\hat{\theta}).$$

Let $R(\theta) \equiv R(\theta, \theta)$. The optimal menu is the solution to

$$\begin{aligned} &\max_{\{q(\cdot),t(\cdot)\}} E_\Theta[t(\Theta) - cq(\Theta)] && \text{(OBJ)} \\ &\text{s.t. } \theta = \arg\max_{\hat{\theta}} R(\theta, \hat{\theta}), && \text{(IC)} \\ &R(\theta) \geq 0, \text{ for every } \theta. && \text{(IR)} \end{aligned}$$

The incentive compatibility (IC) constraint ensures that it is in the best interest of the type-θ retailer to select the quantity-payment pair $(q(\theta), t(\theta))$. The individual rationality (IR) constraint ensures that the retailer accepts the contract because her expected profit by choosing the intended contract is no less than her reservation profit, which without loss of generality, is normalized to zero. We characterize the solution in the following proposition.

Proposition 2. *The optimal menu* $\{q^*(\theta), t^*(\theta)\}$ *is*

$$\begin{aligned} q^*(\theta) &= \mu_0 + \sqrt{1-a^2}\sigma_0\theta + a\sigma_0 z^*(\theta) && (9) \\ t^*(\theta) &= p\left(\mu_0 + \sqrt{1-a^2}\sigma_0\theta - a\sigma_0[\phi(z^*(\theta)) - z^*(\theta)\overline{\Phi}(z^*(\theta))] - \sqrt{1-a^2}\sigma_0 \int_{-\infty}^{\theta} \Phi(z^*(x))dx\right), \end{aligned}$$

where $z^*(\theta)$ *is the unique solution to*

$$a[p\overline{\Phi}(z^*(\theta)) - c] - p\sqrt{1-a^2}\phi(z^*(\theta))\overline{\Phi}(\theta)/\phi(\theta) = 0. \qquad (10)$$

Under the optimal menu, the type-θ *retailer's expected profit is*

$$R(\theta) = p\sqrt{1-a^2}\sigma_0 \int_{-\infty}^{\theta} \Phi(z^*(x))dx, \qquad (11)$$

and the manufacturer's expected profit is

$$M^n = E_\Theta\left[(p-c)\mu_0 - a\sigma_0\Gamma(z^*(\Theta)) - p\sqrt{1-a^2}\sigma_0\Phi(z^*(\Theta))\overline{\Phi}(\Theta)/\phi(\Theta)\right]. \qquad (12)$$

To gain a better understanding of the optimal menu, we first note that $q^*(\theta)$ strictly increases in θ (this follows from Lemma A2 in the Appendix). This is intuitive because it simply says an order quantity intended for an optimistic-forecast-observing retailer is greater than that for a pessimistic-forecast-observing retailer. The monotone property of $q^*(\theta)$ implies the existence of its inverse function, denoted by $\theta^*(q)$, i.e., $\theta^*(q^*(\theta)) = \theta$. Consequently, the optimal menu is equivalent to the payment schedule $T^*(q) \equiv t^*(\theta^*(q))$, which, by simply specifying the transfer payment for any given quantity, is a conceptually simpler way to implement the optimal menu of quantity-payment pairs $\{q^*(\theta), t^*(\theta)\}$.

Under payment schedule $T^*(q)$, $(d/dq)T^*(q)$ can be interpreted as the *marginal wholesale price*, because it is the price the retailer pays for the last unit. A payment schedule in which the marginal wholesale price is decreasing in the quantity purchased is a *quantity-discount scheme*, whereas a schedule in which the marginal wholesale price is increasing in the quantity is a *quantity-premium scheme*. In stochastic-demand settings that are distinct from our own in that, *inter alia*, common information is assumed, Tomlin (2000) and Cachon (2003) show that both quantity-discount schemes and quantity-premium schemes can be effective tools in encouraging efficient quantity decisions to the benefit of individual firms. In principle, in our setting with asymmetric information about demand, it is an open question as to whether the optimal payment scheme exhibits quantity discounts, quantity premia, or a combination of the two.

Proposition 3. (*a*) *The optimal payment schedule* $T^*(q)$ *is a quantity-discount scheme:*

$$(d^2/dq^2)T^*(q) < 0.$$

(*b*) *The marginal wholesale price in the optimal payment scheme,* $(d/dq)T^*(q)$, *is strictly decreasing in the retailer's forecasting accuracy* a.

Quantity discounts are commonly observed in practice, and distinct explanations have been offered for their use. Quantity-discount schemes have been shown to be effective tools in encouraging larger quantity decisions, to the benefit of firms, in settings with stochastic demand (Tomlin 2000 and Cachon 2003) and in settings with deterministic demand but fixed order costs (Weng 1995, Corbett and de Groote 2000, and Chen et al. 2001). A buyer that does not internalize a supplier's fixed order processing cost will order frequently in small batches, and so quantity discounts are a natural mechanism to encourage the buyer to order in a fashion that reflects the supplier's economies of scale. Burnetas et al. (2007) show that quantity discounts can be effective in a setting with asymmetric demand information. Proposition 3a provides a stronger result: Quantity discounts emerge endogenously as an optimal response to private demand-forecast information; see Zhang et al. (2008) for a similar result in a considerably different setting.

To see the intuition as to why quantity discounts are optimal, consider the manufacturer's objectives in offering a contract: differentiating among retailers that have observed different signals, encouraging each to purchase roughly the system-wide-efficient quantity (so as to maximize system profit), and extracting a large a portion of the surplus from the retailer. The intuition for the optimality of quantity discounts is easiest to see when the retailer, after observing her forecast, still faces considerable uncertainty about demand. In this case, her purchase quantity is sensitive to the marginal wholesale price. Virtually all retailers (all but those observing the most pessimistic forecasts), will anticipate being able to sell the first few units they acquire, so the marginal value of these first units will be approximately the retail price. The marginal value of additional units is decreasing, but the extent of this decrease depends on the retailer's privately observed forecast. Charging a high marginal wholesale price for the first units (nearly the retail price) and charging progressively smaller marginal wholesale prices for larger quantities accomplishes two objectives: First, it makes the (low-quantity) contracts intended for retailers that observed unfavorable forecasts unattractive to retailers that have observed favorable forecasts, which limits the profit the favorable-forecast-observing retailer can extract. More generally, making the marginal wholesale price move in tandem with the marginal value of units to the retailer limits the surplus the retailer can capture. Second, it minimizes the quantity distortion for the high-type retailers, which is important because potential system profits are the largest (and hence the impacts of quantity distortions most significant) under favorable forecasts.

Proposition 3b establishes that the impact of the retailer's forecasting accuracy on the optimal unit price extends from the case with the wholesale price contract (see Lemma 1) to the case with the general procurement contract: As the retailer's forecast accuracy worsens (a increases), the *marginal wholesale price* decreases in the optimal procurement contract, just as the *wholesale price* decreases in the optimal wholesale price contract. The basic intuition extends from the wholesale price contract case (see the discussion following Lemma 1). The implication of Proposition 3b is that as the retailer's forecast accuracy deteriorates, the optimal payment schedule $T^*(q)$ "flattens." Figure 1 depicts the

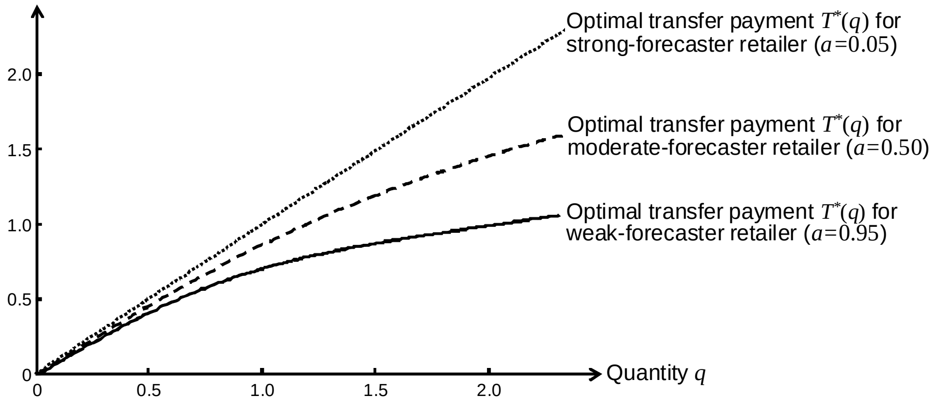

Figure 1: Optimal Procurement Contract for Different Levels of Retailer Forecasting Accuracy. Parameters are $\mu_0 = 1$, $\sigma_0 = 0.5$, $p = 1$, and $c = 0.2$.

optimal procurement contract as a function of the retailer's forecasting accuracy. When the retailer is a weak-forecaster ($a = 0.95$), the optimal contract exhibits substantial quantity discounts, for the reasons described above. In contrast, when the retailer is a strong-forecaster ($a = 0.05$), the optimal contract exhibits little in the way of quantity discounts. In this case, the retailer is quite certain about the actual demand based on her forecast, and so her purchase quantity is rather insensitive to the transfer payment. More specifically, when a is very small, the retailer that has observed forecast θ knows that demand will be very close to the posterior mean $\mu_0 + \sqrt{1-a^2}\sigma_0\theta$; the marginal value of the first $\mu_0 + \sqrt{1-a^2}\sigma_0\theta$ units is approximately the retail price and the marginal value of additional units is roughly zero. Accordingly, a contract which specifies a (constant marginal) wholesale price that is slightly less than the retail price, differentiates among retailers that have observed different forecasts, encourages them to purchase nearly-system-wide-efficient quantities (the retailer that observed forecast θ will purchase almost precisely $\mu_0 + \sqrt{1-a^2}\sigma_0\theta$ units), and allows the manufacturer to appropriate nearly all of the system profit. Thus, a contract where the transfer payment is nearly linear in the quantity purchased is optimal.

As Figure 1 demonstrates, the optimal contract $T^*(q)$ may exhibit significant nonlinearity. If a nonlinear contract is undesirable, the manufacturer will achieve the same profit by instead offering a menu of *linear* contracts (or, equivalently, a menu of two-part tariffs):

$$T^*(q, \theta) \equiv T^*(q^*(\theta)) + (d/dq)T^*(q^*(\theta)) \cdot [q - q^*(\theta)]. \tag{13}$$

The linear contract intended for the type-θ retailer is simply the straight line that is tangent to the concave curve $T^*(q)$ at $q = q^*(\theta)$. In the menu (13), the retailer is given the choice of selecting a contract with a low per-unit price (small $(d/dq)T^*(q^*(\theta))$) and a high fixed payment (large $T^*(q^*(\theta)) - (d/dq)T^*(q^*(\theta)) \cdot q^*(\theta)$), or a contract with a high per-unit price and a low fixed payment.

Before turning to our main focus–the manufacturer's profit under the optimal procurement contract–it is useful to briefly comment on the retailer's profit. As is standard in adverse selection models, the retailer's informational advantage over the manufacturer translates into profit for the retailer (11). The source of this profit is that a type-θ retailer can threaten to select a contract intended for a retailer that has observed a less optimistic forecast, and discouraging the retailer from doing so requires making the contract intended for the type-θ retailer sufficiently attractive to her. The retailer's expected profit is

$$\begin{aligned} E_\Theta[R(\Theta)] &= E_\Theta\left[p\sqrt{1-a^2}\sigma_0 \int_{-\infty}^{\Theta} \Phi(z^*(x))dx\right] \\ &= E_\Theta[p\sqrt{1-a^2}\sigma_0 \Phi(z^*(\Theta))\overline{\Phi}(\Theta)/\phi(\Theta)]. \end{aligned} \tag{14}$$

This quantity is referred to as the retailer's expected information rent.

We now turn to the second and main topic of this section: how the retailer's forecasting accuracy a impacts the manufacturer's expected profit M^n. The manufacturer's expected profit (see equation (12)) is equal to the expected profit from satisfying the mean demand, $(p-c)\mu_0$, minus the expected cost of supply/demand mismatch, $E_\Theta[a\sigma_0\Gamma(z^*(\Theta))]$, and minus the expected information rent captured by the retailer, (14). As the retailer's forecasting accuracy improves, the manufacturer can tailor its contract so that the production quantity reflects this more precise demand information, reducing the expected cost of supply/demand mismatch. On the other hand, as the retailer's informational advantage over the manufacturer increases, it is natural that the information rent captured by the retailer would increase. Whether or not the manufacturer benefits from improved retailer forecasting accuracy depends on the trade-off between the cost of supply/demand mismatch and the retailer's information rent. Because each of these quantities depend on $z^*(\theta)$, to build understanding of the impact of forecasting accuracy a on the manufacturer's profit, we first examine its impact on $z^*(\theta)$. From (9), $z^*(\theta)$ is the number of standard deviations of safety stock purchased by the θ-type retailer in the optimal contract; we refer to $z^*(\theta)$ as the *safety stock factor*.

It is easy to check that $z^*(+\infty) = z^I$; further, $z^*(\theta)$ is strictly increasing in θ (see Lemma A2 in the Appendix). The implication is that only the highest-type retailer orders the system-optimal safety stock and the other types always order less. This is because the manufacturer distorts the quantities downward to limit the information rents earned by the retailer. The result of no distortion for the highest type and downward distortion for other types is a typical result in adverse selection.

Lemma 2. *For every θ, the safety stock factor $z^*(\theta)$ strictly increases in the retailer's forecasting accuracy a.*

In other words, as the retailer's forecasting accuracy improves (a decreases), the manufacturer's optimal contract lowers the safety stock factor for every type retailer. The intuition is as follows. As the retailer's informational advantage grows, she is able to more accurately assess the value of various quantities of units to her; consequently, when faced with a fixed menu of contracts, the retailer is able to make a better-informed contract choice, which increases her information rent. To recapture a portion of the retailer's profit, it is optimal (see Proposition 3b) for the manufacturer to increase the marginal wholesale price $(d/dq)T^*(q)$ across the full range of quantities q. In response, the retailer selects a contract with a smaller safety stock factor. Pushing down the safety stock factor reduces the the retailer's information rent (see (14)). However, at the same time, distorting the safety stock factor downward sacrifices system-wide efficiency by increasing the expected cost of supply/demand mismatch. The larger the retailer's informational advantage, the more willing the manufacturer is to sacrifice system-wide efficiency to ameliorate the information rent paid to the retailer.

We now turn to the paper's main result.

Proposition 4. (a) *The manufacturer's profit under the optimal procurement contract, M^n, is strictly convex in the retailer's forecasting accuracy a; there exists a threshold $a^n \in (0, 1)$ such that M^n is strictly decreasing in a for $a \in (0, a^n)$ and strictly increasing for $a \in (a^n, 1)$.*

(b) *The threshold a^n is solely determined by the ratio c/p (i.e., a^n is independent of all other parameters).*

Figure 2 provides a critical supplement to Proposition 4b: a^n is strictly increasing in c/p.

Interestingly, the convexity result under the optimal wholesale price contract (Proposition 1) also holds under the optimal menu (Proposition 4a).[3] There exists two forecasting regimes: $a \in (0, a^n)$

---

[3] An implication of the convexity result is that the manufacturer's profit is unimodal in any measure of forecasting

and $a \in (a^n, 1)$. Unlike the wholesale price contract case where the two regimes may degenerate into one, here the two regimes always coexist. Although the results under the two contract forms are similar, the explanations are distinct. Here the results are driven by the impact of the retailer's forecasting accuracy on the expected cost of supply/demand mismatch and the expected information rent captured by the retailer. The expected cost of supply/demand mismatch is the product of an exogenous term, $a\sigma_0$, the standard deviation of the posterior demand, and an endogenous term $E_\Theta[\Gamma(z^*(\Theta))]$, which represents the normalized cost of supply/demand mismatch.

First, the manufacturer benefits from improved retailer forecasting accuracy if the retailer is already very good at forecasting, i.e., $a \in (0, a^n)$. The intuition is as follows. When the retailer has strong forecasting capabilities, the optimal contract distorts the safety stock factor downward (Lemma 2), which causes the normalized cost of supply/demand mismatch to be relatively large. Therefore, further improvement in the retailer's forecasting accuracy (reducing a) significantly reduces the total cost of supply/demand mismatch. Because the optimal contract is stingy (characterized by high marginal wholesale prices and low safety stock factors), improved retailer forecasting accuracy has a relatively minor impact on the retailer's information rent. Therefore, the manufacturer benefits from improved retailer forecasting accuracy because the positive impact on reduced supply/demand mismatch cost outweighs any potential negative impact from increased information rents.

Second, the manufacturer is hurt by improved retailer forecasting accuracy if the retailer is very poor at forecasting, i.e., $a \in (a^n, 1)$. The intuition mirrors that of the strong-forecaster case. When the retailer has weak forecasting capabilities, the optimal contract is generous (characterized by low marginal wholesale prices and high safety stock factors). Under a generous contract, increasing the retailer's informational advantage over the manufacturer translates into substantially larger retailer information rent. In contrast, because the safety stock factors are close to the system-wide optimal level, the normalized supply/demand mismatch cost is small, and consequently the savings on the cost of supply/demand mismatch are minor. Consequently, when retailer forecasting accuracy improves, the losses from larger information rent dominate, and the manufacturer is hurt.

The size of the region in which the manufacturer is hurt by selling to a better-forecasting retailer

---

accuracy that is a monotone function of a (or σ_1).

Figure 2: Threshold a^n as Function of the Production Cost to Retail Price Ratio c/p.

depends on the value of the threshold a^n. Under the optimal wholesale price contract, the parallel threshold a^w depends on the product's economics (through the ratio of the production cost to the retail price c/p) and the parameters of the underlying demand distribution (μ_0 and σ_0). In contrast, under the optimal procurement contract, the threshold a^n depends only on the ratio c/p (Proposition 4b). Because this ratio is restricted to a limited range ($c/p \in (0, 1)$), a^n can be completely characterized for all problem parameters in a simple figure, Figure 2. Figure 2 shows that a^n is strictly increasing in c/p.

So, when should a manufacturer be especially concerned that he will be hurt by selling to a better-forecasting retailer? The region in which this outcome occurs is larger (a^n is smaller) when the retail price p is high and the production cost c is low. (This parallels the result under the wholesale price contract.) The intuition is that when the production cost is a small fraction of the the retail price, it is optimal to offer a generous contract so as to encourage the retailer to purchase a large quantity (large safety stock factor). As described immediately above, when the contract is generous, the losses from larger information rent dominate the savings from smaller cost supply/demand mismatch, and so the manufacturer is hurt by improved retailer forecasting.

To summarize, manufacturers that sell high-margin products (e.g., innovative products (e.g., leading-edge electronics), information goods (e.g., books), or goods with strong brands (e.g., Apple, Nike, Polo), where the production cost is small relative to the retail price) tend to be hurt by selling to better-forecasting retailers; this tendency is reinforced if the manufacturer is currently selling to a retailer with weak forecasting capabilities. Manufacturers that sell low-margin products (e.g., mature computer hardware, where the production cost is high relative to the retail price) tend to benefit by selling to better-forecasting retailers; this tendency is reinforced if the manufacturer is currently selling to a retailer with strong forecasting capabilities. These conclusions require some qualification: They apply to products with short life cycles and roughly fixed retail prices, so that the newsvendor framework is appropriate, and they apply when "better-forecasting" is interpreted in the marginal sense.

Proposition 4 also speaks to the setting in which there is a discrete pool of prospective retailers, which requires an understanding that goes beyond the impact of marginal changes in forecasting accuracy. When the manufacturer faces a pool of prospective retailers, the implications of the convexity of the profit function are similar to those discussed in §5: If the pool includes a very strong forecaster, the manufacturer should select the strongest forecaster. If the retailers are all sufficiently weak forecasters, the manufacturer should select the weakest forecaster. All other things being equal (i.e., holding the forecasting accuracy of each retailer fixed), this scenario is more likely to occur when the production cost is small relative to the retail price.

The convexity of the manufacturer's profit function also has implications for the value of demand-forecast information to the manufacturer. In our base setting, the retailer has more information about market demand than the manufacturer. However, in some settings the manufacturer may be able, through its own efforts, to obtain this additional information, eliminating the retailer's informational advantage. This may be the case, for example, when the additional demand information is set of historical sales data (which the manufacturer can piece together by working with its partners, or perhaps directly purchase) or a third-party market demand analysis. Corollary 1 characterizes the value to the manufacturer of acquiring the additional forecast information possessed by the retailer, as a function of the accuracy of that information.

Corollary 1. *There exists* $a^v < a^n$ *such that the manufacturer's gain in expected profit from observing the forecast* Θ *is strictly increasing in* a *for* $a \in (0, a^v)$ *and is strictly decreasing for* $a \in (a^v, 1)$.

Acquiring forecast information to eliminate the retailer's informational advantage is the most valuable when the retailer's informational advantage is moderate. Intuitively, one might expect that the value of eliminating the retailer's informational advantage would be increasing in the size of the informational advantage, because a large informational advantage should translate into large informational rent for the retailer. To see why this conjecture is incorrect, it is useful to consider the extreme case where the retailer's forecasting accuracy is very high (a is very small). In this case, as noted in the discussion following Proposition 3, the optimal contract allows the manufacturer to capture nearly the integrated system profit. Thus, in this extreme case, the manufacturer eliminates the information rents almost completely even though the retailer has superior information. This result suggests that it is not the retailer's superior information that drives the information rents, but rather the extent to which the incentive compatibility (IC) constraint has bite, i.e., the extent to which achieving efficient quantity self-selection requires distorting contracts intended for pessimistic-forecast-observing retailers so that they are unappealing to optimistic-forecast-observing retailers. When the retailer's forecasting accuracy is very high, very little of this distortion is required because a contract with a high marginal wholesale price and small quantity is naturally unappealing to an optimistic-forecasting retailer.

An implication of Corollary 1 is that if forecasting efforts (efforts that reveal forecast Θ) are costly to the manufacturer, it is optimal for the manufacturer to exert these efforts if and only if the retailer's forecasting accuracy is moderate: $a \in (\underline{a}, \overline{a})$, where $\underline{a}$ and $\overline{a}$ satisfy $0 < \underline{a} \leq a^v \leq \overline{a} < 1$.

7 Numerical Study

So far, we have focused on the impact of the retailer's forecasting accuracy on the manufacturer's profit under both a simple and a more complex contractual form. In this section, first, we examine the impact of the retailer's forecasting accuracy on the *retailer's* profit; second, we examine the value to the manufacturer of employing the more complex contractual form.

At the outset, it is unclear whether the retailer benefits from having improved forecasting accuracy. On the one hand, having more accurate demand information allows the retailer to make a better order-quantity decision to alleviate the cost of supply/demand mismatch. On the other hand, a more-accurate-forecasting retailer faces less demand uncertainty, and consequently her purchase quantity is less sensitive to her acquisition cost; consequently, the profit-seeking manufacturer may respond by offering stingier contractual terms. Whether the retailer benefits from improved forecasting accuracy depends on the trade-off between these two factors. To investigate the impact of the retailer's forecasting accuracy on her expected profit, we conducted a numerical study. We fixed $p = 1$ and $\mu_0 = 1$, and varied the other three parameters: $\sigma_0 \in \{0.10, 0.12, 0.14, ..., 0.30\}$, $c \in \{0.20, 0.25, 0.30, ..., 0.80\}$, $a \in \{0.01, 0.03, 0.05, ..., 0.99\}$. Thus, we tested in total 7150 instances. For each instance, we computed the retailer's expected profit under the optimal wholesale price contract and under the optimal procurement contract, i.e., $R^w \equiv E_\theta[R(w^*, \theta)]$ (see (8)) and $R^n \equiv E_\theta[R(\theta)]$ (see (14)). We observed that for every fixed value of σ_0 and c, R^w increases in a and R^n first increases and then decreases in a.

Under the optimal wholesale price contract, the retailer's expected profit R^w increases as her forecasting accuracy worsens. The intuition is that as the retailer's forecasting accuracy worsens, she becomes increasing ill-informed in making a quantity decision. Consequently, the manufacturer, being only able to adjust the unit price, needs to offer a low price to entice the retailer to buy a quantity of any magnitude. Thus, the contractual-terms effect dominates.

Under the optimal procurement contract, the retailer's expected profit R^n increases and then decreases as her forecasting accuracy improves. To see the intuition, it is helpful to consider the extreme cases. As noted previously, when the retailer is a very strong forecaster ($a \approx 0$), the optimal contract achieves nearly the entire integrated system profit for the manufacturer, leaving very little profit for the retailer. When the retailer is a very weak forecaster ($a \approx 1$), the integrated system optimal quantity does not depend on the retailer's privately observed forecast (from (5), $q^I(\theta) \approx \mu_0 + \sigma_0 z^I$ for all θ); consequently, the manufacturer can extract nearly the entire integrated system profit by offering a contract in which the price of purchasing this quantity is the expected revenue it generates, and other quantities are priced sufficiently high as to be made unattractive to the retailer. Consequently, as a strong forecaster's forecast accuracy improves (small a decreases) or a weak forecaster's accuracy worsens (large a increases), the retailer's expected profit decreases toward zero.

A retailer can improve her forecasting accuracy by exerting effort to acquire and process demand-relevant data. The numerical results suggest that the retailer should be wary of improving her forecast accuracy. The retailer should be particularly concerned when she is already a strong forecaster; under such circumstances, even when the cost of improving her forecast accuracy is ignored, improved accuracy results in a loss to the retailer. In contrast, if the retailer is a poor forecaster (lacks an informational advantage over the manufacturer), then whether or not she will benefit from improved forecast accuracy depends on the contractual form she faces: She benefits under the general procurement contract, but is hurt under the wholesale price contract. Taylor (2006) establishes similar results in a much simpler model. Before proceeding, we note that our results rely on the assumption that the manufacturer knows or can infer the retailer's forecasting accuracy. If the retailer is able to improve her accuracy without the manufacturers knowledge, the results may differ.

We now turn to the second topic of this section. Because the general procurement contract is a generalization of the wholesale price contract, the manufacturer's expected profit is higher under the procurement contract: $M^n \geq M^w$. However, the general procurement contract is more complex and will typically be more costly to implement than the simple wholesale price contract. Consequently, in deciding what form of contract to offer, the manufacturer must weigh this implementation cost against the gain from employing the more complex contract $M^n - M^w$. To assess how the exogenous parameters impact the size of this gain, we calculated the gain for every instance of our numerical study. Taking all other parameters as fixed, we observed that in every instance, $M^n - M^w$ increases in a, decreases in c, and first increases and then decreases in σ_0.

The intuition for the impact of the retailer's forecasting accuracy a is most transparent in the extreme case where the retailer is a very strong forecaster ($a \approx 0$). In this case, as noted above, under both contractual forms, the manufacturer's profit is very nearly the entire integrated system profit, so the gain in profit $M^n - M^w$ is small, decreasing as a approaches zero. Similarly, when the production cost c is large, the manufacturer's profit under both contractual forms is small, so the difference in profit $M^n - M^w$ is consequently small. The intuition for the impact of σ_0, the standard deviation of the demand unconditioned on the forecast, is also most easily seen at the extremes. When there is substantial underlying demand uncertainty (σ_0 is large), then the manufacturer's profit under both contractual forms is small, so the difference in profit $M^n - M^w$ is consequently small. When there is little underlying demand uncertainty (σ_0 is small), then the manufacturer can extract nearly the entire integrated system profit by offering a wholesale price contract with a wholesale price very close to the retail price; consequently, the gain in profit from using the more complex contract $M^n - M^w$

is small.

The numerical results suggest that the manufacturer gains the most by employing the more complex contract when the retailer is a poor forecaster (lacks an informational advantage over the manufacturer), the production cost is small, and the underlying demand uncertainty is moderate. If any of these factors is not present and if the incremental administrative cost resulting from using the more complex procurement contract is nontrivial, the manufacturer may be better off by employing the simpler wholesale price contract.

8 Robustness

Although the normal/normal conjugate pair used to model the demand/signal relation has been widely adopted in Bayesian models, the normal distribution assumption is somewhat restrictive. In reality, demand may not be normally distributed; further, the normal assumption may result in a realization of negative demand.[4] In this section, we consider an alternative way of modeling the signal/demand relation, where the normal distribution assumption is relaxed. We show that our central result continues to hold: Improvement in the retailer's forecasting accuracy hurts (benefits) the manufacturer when the retailer is a weak (strong) forecaster.

Consider the following demand model. The market demand faced by the retailer is

$$D = \mu + (1-a)X + aY, \tag{15}$$

where μ and a are constants ($\mu > 0$ and $a \in (0,1)$) and X and Y are mean-zero random variables. Assume that μ is sufficiently large that the probability of D being negative is negligible. Although the distributions of X and Y are common knowledge of both parties, only the retailer observes the realized value X; as before, the retailer observes this signal before choosing her order quantity. We impose the mild assumption that both X and Y have an increasing failure rate.

In the demand model (15), $1-a$ and a are the weights attached to each of the two sources of demand uncertainty, X and Y. Because the retailer observes the realized value of X, the smaller the value of a, the less demand uncertainty faced by the retailer after observing X. Thus, a can be

---

[4] For example, for the threshold $a^w < 1$ it must be that $\sigma_0/\mu_0 > \sqrt{\pi/2} \simeq 1.25$, which implies that the probability that $D < 0$ is non-trivial.

viewed as a measure of the retailer's forecasting accuracy, where, as in the base model, forecasting accuracy improves as a decreases.

Let $\mathbf{M^w}$ ($\mathbf{M^n}$) denote the manufacturer's expected profit under the optimal wholesale price contract (optimal procurement contract) and demand model (15). Proposition 5a establishes that our main result for the wholesale price contract, Proposition 1, extends to this alternative formulation, where the normal distribution assumption is relaxed. Proposition 5b establishes that our main result for the general procurement contract, Proposition 4a, also extends.

Proposition 5. *Suppose demand is given by (15).*

(a) The manufacturer's profit under the optimal wholesale price contract, $\mathbf{M^w}$, *is strictly convex in the retailer's forecasting accuracy* a. *There exists a threshold* $\mathbf{a^w} \in (0,1)$ *such that* $\mathbf{M^w}$ *is strictly decreasing in* a *for* $a \in (0, \mathbf{a^w})$ *and strictly increasing for* $a \in (\mathbf{a^w}, 1)$.

(b) The manufacturer's profit under the optimal procurement contract, $\mathbf{M^n}$, *is strictly convex in* a*; there exists a threshold* $\mathbf{a^n} \in (0,1]$ *such that* $\mathbf{M^n}$ *is strictly decreasing in* a *for* $a \in (0, \mathbf{a^n})$ *and strictly increasing for* $a \in (\mathbf{a^n}, 1)$.

9 Discussion

Our main finding is that a manufacturer's expected profit is convex in the forecasting accuracy of its retail partner. This convexity result lends insight into two managerial questions. First, when faced with a pool of prospective retailers, *ceteris paribus,* should a manufacturer select a retailer that has strong, weak, or intermediate forecasting capabilities? The convexity result implies: If none of the retailers is sufficiently strong, the manufacturer should choose the *weakest* forecaster; otherwise, the manufacturer should choose the strongest forecaster.

Second, does a manufacturer benefit when his retail partner improves her forecasting capabilities? For concreteness, consider a manufacturer partnered with Best Buy, and a manufacturer partnered with Circuit City. As noted previously, historically, Circuit City has been viewed as being a less capable forecaster. In the mid-00's each retailer made significant operational changes which analysts believed would improve the retailer's forecasting capabilities (Wewer and Ma 2004, Ma and Schmitt 2007). Our convexity result implies: A manufacturer benefits by improved forecasting at its retail partner if and only if the retailer is already a good forecaster. To the extent that Best Buy and

Circuit City were quite distinct in their forecasting capabilities, our model predicts that a marginal improvement in forecasting capabilities at these two retailers would have an opposite effect on their manufacturer-partners. Improved forecasting by a strong-forecaster retailer makes a “good” situation better (for the manufacturer), whereas the same improvement by a weak-forecaster makes a “bad” situation worse (for the manufacturer).

We establish that the optimal procurement contract exhibits quantity discounts: Quantity discounts emerge endogenously as an optimal response to the retailer’s private demand-forecast information. Under the optimal contract, the manufacturer tends to be hurt by improved retailer forecasting when the product economics are lucrative. Conversely, the manufacturer tends to benefit by improved retailer forecasting when the product economics are poor. We conclude that a manufacturer should be most concerned about improvements in retailer forecasting accuracy when the retailer is a poor forecaster and the product economics are lucrative.

This work suggests several directions for future research. First, the retailer’s forecasting accuracy, which we have taken as exogenous, could be made endogenous as in Shin and Tunca (2007) and Taylor and Xiao (2007). The assumption of exogenous forecasting accuracy is appropriate when a retailer’s forecasting capabilities are based on long-run decisions that are fixed over the time-scale associated with contracting. A model with endogenous forecasting accuracy would capture the setting where a retailer can change her forecasting accuracy over a shorter time-scale. Second, the space of contracts could be enlarged to include contracts where transfer payments are based not only on the quantity purchased, but also on the retailer’s realized sales, as in returns or rebates contracts. Third, other types of asymmetric information could be explored. Private information abounds in operational settings and often even better-informed firms have imperfect information. For example, manufacturers tend to have private, but imperfect, information about their production costs and retailers tend to have similar information about their shelf-space costs. Our convexity result suggests that firms prefer to partner with firms that have either quite accurate or quite inaccurate information. Further research is required to see if this result carries over to operational settings with private information about factors other than demand.

Appendix

Lemma A1 is useful in the proof of Proposition 1. Let $G(z) \equiv (p\overline{\Phi}(z) - c)(\mu_0 + a\sigma_0 z)$.

Lemma A1. $G(z)$ is unimodal for $z \leq \overline{\Phi}^{-1}(c/p)$.

Proof of Lemma A1. Note that

$$\begin{aligned} G'(z) &= (p\overline{\Phi}(z) - c)a\sigma_0 - p\phi(z)(\mu_0 + a\sigma_0 z) \\ &= \phi(z)[a\sigma_0 H(z) - p\mu_0], \end{aligned}$$

where $H(z) \equiv p\overline{\Phi}(z)/\phi(z) - c/\phi(z) - pz$. To prove that $G(z)$ is unimodal, it suffices to show that $G'(z)$ changes the sign at most once for $z \leq \overline{\Phi}^{-1}(c/p)$; to do so, it is sufficient to show that $H'(z) < 0$ for $z \leq \overline{\Phi}^{-1}(c/p)$. Because $\overline{\Phi}(z)/\phi(z)$, $-c/\phi(z)$, and $-pz$ strictly decrease for $z \geq 0$, $H'(z) < 0$ for $z \geq 0$. Now suppose $z < 0$. Then

$$\begin{aligned} H'(z) &= p\frac{-\phi^2(z) + z\phi(z)\overline{\Phi}(z)}{\phi^2(z)} - c\frac{z\phi(z)}{\phi^2(z)} - p \\ &= -2p + \frac{z[p\overline{\Phi}(z) - c]}{\phi(z)} \\ &< 0, \end{aligned}$$

where the last inequality follows from $p\overline{\Phi}(z) - c \geq 0$ (because $z \leq \overline{\Phi}^{-1}(c/p)$) and $z < 0$. ■

Proof of Proposition 1. To prove the first part of the proposition, it suffices to show that $(d^2/da^2)M^w > 0$. Because $M^w = G(z^*)$, by the Envelope Theorem,

$$\frac{dM^w}{da} = (p\overline{\Phi}(z^*) - c)\sigma_0 z^*,$$

which leads to

$$\frac{d^2M^w}{da^2} = \sigma_0(p\overline{\Phi}(z^*) - c - pz^*\phi(z^*))\frac{dz^*}{da}.$$

Because z^* is the maximizer of $G(z)$ over $z \leq \overline{\Phi}^{-1}(c/p)$, the first-order necessary condition is $G'(z^*) \geq 0$, i.e., $(p\overline{\Phi}(z^*) - c)a\sigma_0 - p\phi(z^*)(\mu_0 + a\sigma_0 z^*) \geq 0$, or equivalently, $[p\overline{\Phi}(z^*) - c - pz^*\phi(z^*)]a\sigma_0 \geq p\phi(z^*)\mu_0 > 0$. This, together with Lemma 1 that $(d/da)z^* > 0$, establishes that $(d^2/da^2)M^w > 0$.

Case 1. $c \geq p/2$. Note that $\overline{\Phi}^{-1}(c/p) \leq 0$. Consider any $a \in (0,1)$. Because $z^* \leq \overline{\Phi}^{-1}(c/p)$, $p\overline{\Phi}(z^*) - c \geq 0$ and $z^* \leq 0$. Therefore, $(d/da)M^w \leq 0$ for any $a \in (0,1)$.

Case 2. $c < p/2$ and $a^w > 1$. Then $p\phi(0)\mu_0 > (p/2 - c)\sigma_0$. Take any $a \in (0,1)$. Note that $G'(0) = (p/2 - c)a\sigma_0 - p\phi(0)\mu_0 < 0$. This, together with Lemma A1, implies that $z^* < 0$. Consequently, $p\overline{\Phi}(z^*) - c > p/2 - c > 0$. Therefore, $(d/da)M^w < 0$ for any $a \in (0,1)$.

Case 3. $c < p/2$ and $a^w < 1$. It follows from the same arguments as in Case 2 that $(d/da)M^w < 0$ for any $a \in (0, a^w)$. Take any $a \in (a^w, 1)$. Note that $G'(0) = (p/2 - c)a\sigma_0 - p\phi(0)\mu_0 > 0$. This, together with Lemma A1, implies that $z^* > 0$. Note that $G'(\overline{\Phi}^{-1}(c/p)) = -p\phi(\overline{\Phi}^{-1}(c/p))(\mu_0 + a\sigma_0\overline{\Phi}^{-1}(c/p)) < 0$ because $\overline{\Phi}^{-1}(c/p) > 0$. Hence $z^* < \overline{\Phi}^{-1}(c/p)$, i.e., $p\overline{\Phi}(z^*) - c > 0$. Therefore, $(d/da)M^w > 0$ for any $a \in (a^w, 1)$. ■

Lemma A2 is useful in the proofs of Lemma 2 and Propositions 2, 3 and 4.

Lemma A2. *$z^*(\theta)$, the solution to (10), is unique. $z^*(\theta)$ strictly increases in θ, a and p and strictly decreases in c.*

Proof of Lemma A2. By dividing the both sides of (10) by $ap\overline{\Phi}(z^*(\theta))$, we have

$$1 - \frac{c}{p\overline{\Phi}(z^*(\theta))} - \frac{\sqrt{1-a^2}}{a}\frac{\phi(z^*(\theta))}{\overline{\Phi}(z^*(\theta))}\frac{\overline{\Phi}(\theta)}{\phi(\theta)} = 0. \tag{16}$$

Let $A(z) \equiv \frac{c}{p\overline{\Phi}(z)} + \frac{\sqrt{1-a^2}}{a}\frac{\phi(z)}{\overline{\Phi}(z)}\frac{\overline{\Phi}(\theta)}{\phi(\theta)}$. Clearly, $A(z)$ strictly increases in z. Because $A(-\infty) = c/p < 1$ and $A(+\infty) = +\infty > 1$, there exists a unique solution $z^*(\theta)$ that satisfies (16) for each θ. Because $A(z)$ strictly decreases in θ, a and p (strictly increases in c), the solution $z^*(\theta)$ strictly increases in θ, a and p and strictly decreases in c. ■

Proof of Proposition 2. The proof proceeds as follows. The bulk of the proof is devoted to identifying a solution to the relaxed contract design problem in which the (IC) constraint is replaced by the corresponding first order necessary condition. We then observe that the solution to this relaxed problem satisfies the constraints of the original problem. Using (3), we can write

$$\begin{aligned}
R(\theta,\hat{\theta}) &= pE_{D_\theta}\min(D_\theta, q(\hat{\theta})) - t(\hat{\theta}) \\
&= p(\mu_0 + \sqrt{1-a^2}\sigma_0\theta) + pa\sigma_0 E_X \min\left\{X, \frac{q(\hat{\theta}) - \mu_0 - \sqrt{1-a^2}\sigma_0\theta}{a\sigma_0}\right\} - t(\hat{\theta}) \\
&= p(\mu_0 + \sqrt{1-a^2}\sigma_0\theta) \\
&\quad -pa\sigma_0\left[\phi(z(\hat{\theta}) + \frac{\sqrt{1-a^2}}{a}(\hat{\theta}-\theta)) - [z(\hat{\theta}) + \frac{\sqrt{1-a^2}}{a}(\hat{\theta}-\theta)]\overline{\Phi}(z(\hat{\theta}) \right. \\
&\quad \left. + \frac{\sqrt{1-a^2}}{a}(\hat{\theta}-\theta))\right] - t(\hat{\theta}),
\end{aligned} \tag{17}$$

where $z(\theta) \equiv [q(\theta) - \mu_0 - \sqrt{1-a^2}\sigma_0\theta]/(a\sigma_0)$ and the last equality follows from (3). It follows from

(IC) and the Envelope Theorem that

$$\begin{aligned} R'(\theta) &= \left.\frac{\partial R(\theta,\hat{\theta})}{\partial \theta}\right|_{\hat{\theta}=\theta} \\ &= p\sqrt{1-a^2}\sigma_0\Phi(z(\theta)) \text{ (by (17))}, \end{aligned}$$

which by integration, leads to $R(\theta) = R(-\infty) + p\sqrt{1-a^2}\sigma_0 \int_{-\infty}^{\theta} \Phi(z(x))dx$. Note that $R(\theta)$ increases in θ. Clearly, at the optimal solution, $R(-\infty) = 0$. Hence we have

$$R(\theta) = p\sqrt{1-a^2}\sigma_0 \int_{-\infty}^{\theta} \Phi(z(x))dx. \tag{18}$$

By definition of $R(\theta)$,

$$\begin{aligned} R(\theta) &= R(\theta,\theta) \\ &= p(\mu_0 + \sqrt{1-a^2}\sigma_0\theta) - pa\sigma_0[\phi(z(\theta)) - z(\theta)\overline{\Phi}(z(\theta))] - t(\theta). \end{aligned} \tag{19}$$

From (18) and (19), we can express $t(\theta)$ by using $z(\cdot)$ as follows:

$$t(\theta) = p(\mu_0 + \sqrt{1-a^2}\sigma_0\theta) - pa\sigma_0[\phi(z(\theta)) - z(\theta)\overline{\Phi}(z(\theta))] - p\sqrt{1-a^2}\sigma_0 \int_{-\infty}^{\theta} \Phi(z(x))dx. \tag{20}$$

Substituting $t(\theta)$ in (OBJ) with the right hand side of the above equation, (OBJ) can be rewritten as a function of $z(\cdot)$:

$$\begin{aligned} &E_\Theta[t(\Theta) - c(\mu_0 + \sqrt{1-a^2}\sigma_0\Theta + a\sigma_0 z(\Theta))] \\ &\quad = E_\Theta[(p-c)(\mu_0 + \sqrt{1-a^2}\sigma_0\Theta) - a\sigma_0\Gamma(z(\Theta)) - p\sqrt{1-a^2}\sigma_0\Phi(z(\Theta))\overline{\Phi}(\Theta)/\phi(\Theta)], \end{aligned}$$

which by pointwise optimization, maximized at $z^*(\theta)$ for every θ where $z^*(\theta)$ is uniquely (see Lemma A2) determined from (10). The corresponding $t^*(\theta)$ can then be determined from (20).

Clearly, the solution $(z^*(\theta), t^*(\theta))$ constructed above yields an upper bound on the manufacturer's expected profit, and also satisfies (IR) and the first-order necessary condition of (IC). From Lemma A2, $z^*(\theta)$ increases in θ, which is a sufficient condition to ensure that the solution $(z^*(\theta), t^*(\theta))$ satisfies (IC). Therefore, $(z^*(\theta), t^*(\theta))$ solves the manufacturer's contract design problem (OBJ)-(IR).■

Proof of Proposition 3. (a) Note that

$$\frac{dt^*(\theta)}{d\theta} = p\sigma_0\overline{\Phi}(z(\theta))[\sqrt{1-a^2} + az'(\theta)] \tag{21}$$

and

$$\frac{dq^*(\theta)}{d\theta} = \sigma_0[\sqrt{1-a^2} + az'(\theta)]. \tag{22}$$

By definition of $T^*(q)$,

$$\begin{aligned}\frac{dT^*(q)}{dq} &= \frac{dt^*(\theta)}{d\theta}\Big/\frac{dq^*(\theta)}{d\theta} \\ &= \frac{p\sigma_0\overline{\Phi}(z(\theta))[\sqrt{1-a^2} + az'(\theta)]}{\sigma_0[\sqrt{1-a^2} + az'(\theta)]} \text{ (by (21) and (22))} \\ &= p\overline{\Phi}(z(\theta)), \end{aligned} \tag{23}$$

where θ is such that $q = q^*(\theta)$. We can also write $(d/dq)T^*(q) = p\overline{\Phi}(Z(q))$, where $Z(q)$ is the z value corresponding to q. Clearly, $Z(q)$ is strictly increasing q. Hence, $(d^2/dq^2)T^*(q) = -p\phi(Z(q))Z'(q) < 0$. (b) It follows from (23) that $(\partial^2/\partial q\partial a)T^*(q) = -p\phi(z(\theta))(\partial/\partial a)z(\theta)$, where θ is such that $q = q^*(\theta)$. Because $(\partial/\partial a)z(\theta) > 0$ (from Lemma A2), $(\partial^2/\partial q\partial a)T^*(q) < 0$.■

Proof of Lemma 2. See Lemma A2.■

Proof of Proposition 4. (a) First we establish the convexity property. By the Envelope Theorem,

$$\frac{dM^n}{da} = E_\Theta\left[-\sigma_0\Gamma(z^*(\Theta)) + \frac{pa\sigma_0}{\sqrt{1-a^2}}\Phi(z^*(\Theta))\frac{\overline{\Phi}(\Theta)}{\phi(\Theta)}\right]. \tag{24}$$

Note that $\Gamma'(z) = -p\overline{\Phi}(z) + c$. It follows from (10) that $p\overline{\Phi}(z^*(\theta)) - c > 0$ for every θ. Thus

$$\Gamma'(z^*(\theta)) < 0,$$

which together with the fact that $z^*(\theta)$ strictly increases in a (see Lemma A2), implies that $\Gamma(z^*(\theta))$ strictly decreases in a. Note that the second part of $(d/da)M^n$ is clearly strictly increasing in a. Consequently, $(d/da)M^n$ strictly increases in a; therefore, M^n is strictly convex in a.

To show the existence of $a^n \in (0,1)$, it suffices to show that $(d/da)M^n|_{a\to 1^-} > 0$ and $(d/da)M^n|_{a\to 0^+} < 0$. It follows from (10) that as $a \to 1^-$, $z^*(\theta) \to z^I$ for every θ. Therefore, by (24), we have $(d/da)M^n|_{a\to 1^-} = +\infty > 0$. Similarly, as $a \to 0^+$, $z^*(\theta) \to -\infty$ for every θ. Therefore, by (24), we have $(d/da)M^n|_{a\to 0^+} = -\infty < 0$.

(b) From (24), the definition of $\Gamma(\cdot)$, and the fact that a^n is the unique solution to $(d/da)M^n = 0$, a^n is the unique solution to

$$E_\Theta \left[-\phi(z^*(\Theta)) + z^*(\Theta)\overline{\Phi}(z^*(\Theta)) - \frac{c}{p} z^*(\Theta) + \frac{a}{\sqrt{1-a^2}} \Phi(z^*(\Theta)) \frac{\overline{\Phi}(\Theta)}{\phi(\Theta)} \right] = 0. \tag{25}$$

From (10), for every θ, $z^*(\theta)$ depends solely on the ratio c/p and a (i.e., $z^*(\theta)$ is independent of all other parameters). Therefore, in terms of the model primitives, the left hand side of (25) depends only on c/p and a. Because for any fixed c/p, the solution to (25), a^n, is unique, a^n is determined solely by the ratio c/p (i.e., a^n is independent of all other parameters).■

Proof of Corollary 1. The manufacturer that observes forecast θ achieves the integrated system profit by offering a contract that induces the retailer to purchase the integrated-system optimal quantity $q^I(\theta)$ for a transfer payment equal to the expected revenue generated by these units under demand D_θ, i.e., $pE_{D_\theta}[\min(D_\theta, q)] = p(\mu_0 + \sqrt{1-a^2}\sigma_0\theta - a\sigma_0[\phi(z^I) - z^I\overline{\Phi}(z^I)])$. An example of a contract that achieves this is the payment schedule $T(q) = pE_{D_\theta}[\min(D_\theta, q)]$ for $q \le q^I(\theta)$ and $T(q) = \infty$ for $q > q^I(\theta)$. The manufacturer's gain in expected profit from observing the forecast is $\Pi^I - M^n$. Because $M^n \le \Pi^I$, $\lim_{a \to 0^+} M^n = \Pi^I$, M^n is strictly convex in a (from Proposition 4), and $(d/da)\Pi^I = -\sigma_0\Gamma(z^I), \lim_{a\to 0^+}(d/da)M^n < -\sigma_0\Gamma(z^I)$. This together with the facts that $(d/da)M^n|_{a=a^n} = 0$ and M^n is strictly convex and continuous in a, implies the existence of a unique $a^v \in (0, a^n)$ satisfying $(d/da)M^n|_{a=a^v} = -\sigma_0\Gamma(z^I)$. Therefore, $(d/da)[\Pi^I - M^n] > 0$ for $a \in (0, a^v)$ and $(d/da)[\Pi^I - M^n] < 0$ for $a \in (a^v, 1)$.■

References

Arya, A., B. Mittendorf. 2004. Using returns policies to elicit retailer information. *RAND J. Econom.* **35**(3) 617-630.

Aviv, Y. 2001. The effect of collaborative forecasting on supply chain performance. *Management Sci.* **47**(10) 1326-1343.

Burnetas, A., S. M. Gilbert., C. E. Smith. 2007. Quantity discounts in single-period supply contracts with asymmetric demand information. *IIE Trans.* **39**(5) 465-479.

Cachon, G. and M. Fisher. 2000. Supply chain inventory management and the value of shared information. *Management Sci.* **46**(8) 1032-48.

Cachon, G. P., 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. 2003. Supply chain coordination with contracts. S. Graves and T. De Kok, eds. *Handbooks in Operations Research and Management Science: Supply Chain Management.* North Holland, Amsterdam.

Chen, F. 2003. Information sharing and supply chain coordination. S. Graves and T. De Kok, eds. *Handbooks in Operations Research and Management Science: Supply Chain Management.* North Holland, Amsterdam.

Chen, F., A. Federgruen, Y.-S. Zheng. 2001. Coordination mechanisms for a distribution system with one supplier and multiple retailers. *Management Sci.* **47**(5) 693-708.

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

Feldman, J., J. Cramer. 2004. Hardlines retail. Report (April 26), SG Cowen, New York.

Fliedner, G., 2003. CPFR: An emerging supply chain tool. *Industrial Management & Data Systems* **103**(1) 14-21.

Lariviere, M.A., E. Porteus. 2001. Selling to the newsvendor: An analysis of price-only contracts. *Manufacturing Service Oper. Management* **3**(4) 293-305.

Lewis, T., D. Sappington. 1991. All-or-nothing information control. *Economics Letters* **37**(2), 111–113.

Lee, H. L., K. C. So, C. S. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Sci.* **46**(5) 626-643.

Ma, V., J. Schmitt. 2007. Circuit City Stores, Inc.: Increasing earnings leverage potential, turnaround accelerating. Report (Feb. 9), CIBC World Markets, New York.

Miyaoka, J., W. H. Hausman. 2008. How improved forecasts can degrade decentralized supply chains. *Manufacturing Service Oper. Management* **10**(3), 547–562.

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

Porteus, E. L. 2002. *Foundations of Stochastic Inventory Theory.* Stanford University Press, Stanford, California.

Qi, F., K. Zhu. 2006. Endogenous information acquisition in supply chain management. Working paper, Hong Kong University of Science and Technology.

Rajan, M., R. Saouma. 2006. Optimal information asymmetry. *The Accounting Review* **81**(3), 677–712.

Ren, Z. J., M. Cohen, T. Ho, C. Terwiesch. 2008. Sharing forecast information in a long-term supply chain relationship: The role of customer review strategy. *Oper. Res.*, forthcoming.

Shin, H., T. Tunca. 2007. Beating the accuracy trap: Overinvestment in demand forecasting and supply chain coordination under downstream competition. Working paper, Stanford University.

Taylor, T. A. 2006. Sale timing in a supply chain: When to sell to the retailer. *Manufacturing Service Oper. Management* **8**(1) 23-42.

Taylor, T. A., W. Xiao. 2007. Incentives for retailer forecasting: Rebates versus returns. Working paper, University of California, Berkeley.

Taylor, T. A., W. Xiao. 2008. Technical appendix to Does a manufacturer benefit from selling to a better-forecasting retailer? Working paper, University of California, Berkeley.

Tomlin, B. 2003. Capacity investments in supply chains: Sharing-the-gain rather than sharing-the-pain. *Manufacturing Service Oper. Management* **5**(4) 317-333.

Weng, Z. K. 1995. Channel coordination and quantity discounts. *Management Sci.* **41**(9) 1509-1521.

Wewer, D., V. Ma. 2004. Best Buy Co.: 2H04 and 2005 outlook more promising. Report (Sept. 16), CIBC World Markets, New York.

Widlitz, S. 2005. CC: SWAT teams and test catalog may be new catalysts – neutral. Report (July 21), Fulcrum Global Partners, New York.

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

Zhang, H., M. Nagarajan, G. Sosic. 2008. Optimal dynamic supplier contracts under asymmetric inventory information. Working paper, University of Southern California.

Internet Appendix

This internet appendix includes the analysis and proofs for the demand model in §8. First, we introduce some notation. Second, we consider the wholesale price contract and prove the result for this case, Proposition 5a. Third, we consider the general procurement contract and prove the result for that case, Proposition 5b.

Let $F_i(\cdot)$ and $f_i(\cdot)$ be the distribution and density of i, for $i = X, Y$. Let $\overline{F}_i(\cdot) \equiv 1 - F_i(\cdot)$. Because X and Y have an increasing failure rate, i.e., $f_i(x)/\overline{F}_i(x)$ strictly increases in x, for $i = X, Y$.

The analysis for the wholesale price contract is similar to that in §5. Given wholesale price w and forecast $X = x$, the retailer's optimal order quantity is

$$q(w, x) = \mu + (1 - a)x + az,$$

where $z = \overline{F}_Y^{-1}(w/p)$. The manufacturer's wholesale price contract design problem is

$$\max_{w \geq c} (w - c) E_X q(w, X),$$

which can be rewritten as

$$\max_{z \leq \overline{F}_Y^{-1}(c/p)} (p\overline{F}_Y(z) - c)(\mu + az).$$

The manufacturer's expected profit under the optimal wholesale price contract, $\mathbf{M^w}$, is the objective value that results from this optimization problem.

Proof of Proposition 5a. Let $H(z) \equiv (p\overline{F}_Y(z) - c)(\mu + az)$. Then

$$\begin{aligned} H'(z) &= (p\overline{F}_Y(z) - c)a - pf_Y(z)(\mu + az) \\ &= \overline{F}_Y(z) \left[pa - \frac{ca}{\overline{F}_Y(z)} - \frac{pf_Y(z)(\mu + az)}{\overline{F}_Y(z)} \right], \end{aligned}$$

which changes sign exactly once because Y has an increasing hazard rate. Therefore, $H(z)$ is unimodal and the maximizer, denoted by z^*, satisfies the first-order condition, i.e.,

$$(p\overline{F}_Y(z^*) - c)a - pf_Y(z^*)(\mu + az^*) = 0.$$

This implies that

$$p\overline{F}_Y(z^*) - c - pf_Y(z^*)z^* > 0. \tag{26}$$

Furthermore, it follows from the Implicit Function Theorem that

$$\frac{dz^*}{da} > 0. \tag{27}$$

By the Envelope Theorem,

$$\frac{d\mathbf{M^w}}{da} = (p\overline{F}_Y(z^*) - c)z^*$$

and hence

$$\begin{aligned} \frac{d^2\mathbf{M^w}}{da^2} &= [p\overline{F}_Y(z^*) - c - pf_Y(z^*)z^*]\frac{dz^*}{da} \\ &> 0, \end{aligned}$$

where the inequality follows from (26) and (27). This proves the convexity. The existence of the threshold $\mathbf{a^w}$ follows from the fact that $(d/da)\mathbf{M^w}|_{a\to 0^+} = -\infty$ and $(d/da)\mathbf{M^w}|_{a\to 1^-} > 0.$■

We now proceed to analyze the general procurement contract. The analysis is similar to that in §6. Consider the menu of quantity-payment contracts $\{q(x), t(x)\}$, where it is in the best interest of the type-x retailer (who observed $X = x$) to order $q(x)$ and pay $t(x)$ to the manufacturer. If the type-x retailer chooses the contract $(q(\hat{x}), t(\hat{x}))$, then her expected profit is

$$R(x, \hat{x}) = pE_Y\{\min(\mu + (1-a)x + aY, q(\hat{x})\} - t(\hat{x}).$$

Let $R(x) \equiv R(x, x)$. The optimal menu is the solution to

$$\max_{\{q(\cdot),t(\cdot)\}} E_x[t(x) - cq(x)] \tag{OBJ}$$

$$\text{s.t. } x = \arg\max_{\hat{x}} R(x, \hat{x}), \tag{IC}$$

$$R(x) \geq 0, \text{ for every } x. \tag{IR}$$

Lemma A3. *The optimal menu* $\{q^*(x), t^*(x)\}$ *is*

$$\begin{aligned} q^*(x) &= \mu + (1-a)x + az^*(x) \\ t^*(x) &= p[\mu + (1-a)x] + paE_Y \min(Y, z^*(x)) - p(1-a)\int_{-\infty}^{x} F_Y(z^*(s))ds, \end{aligned}$$

where $z^*(x)$ *is the unique solution to*

$$a[p\overline{F}_Y(z^*(x)) - c] - p(1-a)f_Y(z^*(x))\overline{F}_X(x)/f_X(x) = 0. \tag{28}$$

Under the optimal menu, the manufacturer's expected profit is

$$\mathbf{M^n} = E_X\left[(p-c)\mu + paE_Y\min(Y, z^*(X)) - caz^*(X) - p(1-a)F_Y(z^*(X))\overline{F}_X(X)/f_X(X)\right]. \quad (29)$$

Proof of Lemma A3. The proof proceeds as follows. The bulk of the proof is devoted to identifying a solution to the relaxed contract design problem in which the (IC) constraint is replaced by the corresponding first order necessary condition. We then observe that the solution to this relaxed problem satisfies the constraints of the original problem. Note that

$$\begin{aligned} R(x,\hat{x}) &= pE_Y[\min\{\mu + (1-a)x + aY, q(\hat{x})\}] - t(\hat{x}) \\ &= p[\mu + (1-a)x] + pa\left[\int_{-\infty}^{\frac{1-a}{a}(\hat{x}-x)+z(\hat{x})} yf_Y(y)dy\right. \\ &\quad \left. + \left[\frac{1-a}{a}(\hat{x}-x) + z(\hat{x})\right]\overline{F}_Y\left(\frac{1-a}{a}(\hat{x}-x) + z(\hat{x})\right)\right] - t(\hat{x}) \qquad (30) \end{aligned}$$

where $z(x) \equiv [q(x) - \mu - (1-a)x]/a$. It follows from (IC) and the Envelope Theorem

$$\begin{aligned} R'(x) &= \frac{\partial R(x,\hat{x})}{\partial x}|_{\hat{x}=x} \\ &= p(1-a)F_Y(z(x)) \text{ (by (30))}, \end{aligned}$$

which by integration, leads to $R(x) = R(-\infty) + p(1-a)\int_{-\infty}^{x} F_Y(z(s))ds$. Note that $R(x)$ increases in x. Clearly, at the optimal solution, $R(-\infty) = 0$. Hence we have

$$R(x) = p(1-a)\int_{-\infty}^{x} F_Y(z(s))ds. \qquad (31)$$

By definition of $R(x)$,

$$\begin{aligned} R(x) &= R(x,x) \\ &= p[\mu + (1-a)x] + pa\left[\int_{-\infty}^{z(x)} yf_Y(y)dy + z(x)\overline{F}_Y(z(x))\right] - t(x). \qquad (32) \end{aligned}$$

From (31) and (32), we can express $t(x)$ by using $z(\cdot)$ as follows:

$$t(x) = p[\mu + (1-a)x] + pa\left[\int_{-\infty}^{z(x)} yf_Y(y)dy + z(x)\overline{F}_Y(z(x))\right] - p(1-a)\int_{-\infty}^{x} F_Y(z(s))ds. \qquad (33)$$

Substituting $t(x)$ in (OBJ) with the right hand side of the above equation, (OBJ) can be rewritten as a function of $z(\cdot)$:

$$\begin{aligned} &E_X[t(X) - c(\mu + (1-a)X + az(X))] \\ &\quad = E_X[(p-c)(\mu + (1-a)X) - caz(X) + pa\left[\int_{-\infty}^{z(X)} y f_Y(y)dy + z(X)\overline{F}_Y(z(X))\right] \\ &\qquad - p(1-a)F_Y(z(X))\overline{F}_X(X)/f_X(X)]. \end{aligned} \tag{34}$$

Before characterizing the $z(x)$ which maximizes (34), we note that, by argument similar to that in the proof of Lemma A2, because Y has an increasing hazard rate, the solution to (28), $z^*(x)$, is unique. By pointwise optimization, (34) is maximized at $z^*(x)$ for every x. The corresponding $t^*(x)$ can then be determined from (33).

Clearly, the solution $(z^*(x), t^*(x))$ constructed above yields an upper bound on the manufacturer's expected profit, and also satisfies (IR) and the first-order necessary condition of (IC). It can be verified from (28) that $z^*(x)$ increases in x (because X and Y are IFR), which is a sufficient condition to ensure that the solution $(z^*(x), t^*(x))$ satisfies (IC). Therefore, $(z^*(x), t^*(x))$ solves the manufacturer's contract design problem (OBJ)-(IR).■

Proof of Proposition 5b. To prove the convexity, note that

$$\begin{aligned} \frac{d\mathbf{M^n}}{da} &= E_X\left[pE_Y \min(Y, z^*(X)) - cz^*(X) + pF_Y(z^*(X))\overline{F}_X(X)/f_X(X)\right] \\ &= E_X\left[p\left[\int_{-\infty}^{z^*(X)} y f_Y(y)dy + z^*(X)\overline{F}_Y(z^*(X))\right] - cz^*(X) + pF_Y(z^*(X))\overline{F}_X(X)/f_X(X)\right] \end{aligned} \tag{35}$$

Hence,

$$\frac{d^2\mathbf{M^n}}{da^2} = E_X\left[\left[p\overline{F}_Y(z^*(X)) - c + pf_Y(z^*(X))\overline{F}_X(X)/f_X(X)\right]\frac{dz^*(X)}{da}\right].$$

It follows from (28) that $p\overline{F}(z^*(x)) - c > 0$ and $(d/da)z^*(x) > 0$ for every x. Therefore, $(d^2/da^2)\mathbf{M^n} > 0$, i.e., $\mathbf{M^n}$ is strictly convex in a.

To show the existence of $\mathbf{a^n} \in (0, 1]$, it suffices to show that $(d/da)\mathbf{M^n}|_{a\to 0^+} < 0$. As $a \to 0^+$, $z^*(x) \to -\infty$ for every x. Therefore, by (35), we have $(d/da)\mathbf{M^n}|_{a\to 0^+} = -\infty < 0$.■