Inducing Forecast Revelation through Restricted Returns

Martin A. Lariviere
Kellogg School of Management
Northwestern University
Evanston IL, 60208

June 17, 2002

Abstract

We consider a supply chain consisting of one supplier selling through one retailer who faces a newsvendor problem. There is a positive probability that the retailer is capable of gaining improved demand information through costly forecasting. The supplier would like to induce the retailer to forecast and share that information. Restricting the retailer's ability to return unsold product would intuitively appear to be a viable way by which to provide the desired incentives. However, it is well known that a generous returns policy increases the supplier's profit. We explore this tension between providing incentives to forecast and capturing channel profits. We examine both price-based returns mechanisms (buy backs) and quantity-based returns mechanisms (quantity flexibility contracts). Thus, a second goal is to compare the relative performance of these two schemes.

We show that under either contract, inducing forecasting retailers to take a different contract from non-forecasting retailers requires restricting returns. The forecaster is charged a lower price than the non-forecaster but has less flexibility in returning product. Using buy backs, the supplier must sacrifice some channel profit to differentiate between forecasting and non-forecasting retailers. With quantity flexibility contracts, reducing channel efficiency is not required to distinguish between the types of retailers. It is consequently some what surprising that buy backs generally result in greater supplier profit than quantity flexibility contracts unless forecasting is very expensive.

Key words: Supply-chain contracting; forecasting; buy backs; quantity flexibility; Bayesian inventory.

I am thankful for the helpful comments of seminar participants at the Kenan-Flagler Business School, The Wharton School, the Department of Industrial and Systems Engineering at Lehigh University, and Columbia Business School.

1 Introduction

Precise information facilitates efficiently matching supply and demand. Forecasting is consequently a critical activity in many supply chains. When a supply chain consists of independent parties, some or all supply-chain members may not gather as much information as a centralized system. Even if all parties collect information, individuals may not reveal all that they have learned. Altering the terms governing supply-chain transactions is a possible way to induce supply-chain members to both gather and share information.

Consider the approach IBM took when launching of a generation of laptop computers (Zarley, 1994). Hoping to see "real" demand, Big Blue offered several resellers a special allocation of laptops if they accepted restricted contract terms. Ordinarily, resellers had significant freedom to cancel or return orders, but the terms of the special allocation imposed "abusive" penalties on returns or cancellation. (Machines could still be purchased under the usual terms.) In the words of IBM's vice president of channel management, the program was intended "to help the channel better forecast demand and to more effectively manage inventory." Restricting returns would hopefully get people "out of game playing" according to an executive at one of the resellers (Zarley, 1994).

IBM's scheme has an intuitive appeal. Limited returns clearly provides an incentive to forecast accurately, but it also begs the question why IBM offered such easy returns in the first place. While there are undoubtedly institutional reasons why generous returns policies have long been standard in the computer industry, there is another important consideration. A well-designed returns policy can greatly improve the performance of a supply chain for a short life cycle product. Pasternack (1985) shows that the greater the flexibility that the upstream player offers the downstream player, the greater the former's profit. Providing incentives to forecast is thus at odds with claiming a large share of supply-chain profit.

We examine this trade off in a setting similar to Pasternack (1985). An upstream supplier sells to a downstream retailer facing a newsvendor problem. The parties agree on an initial estimate of the demand distribution. If forecasting takes place, information is observed and beliefs about demand are updated. Only the retailer can forecast, and he incurs some fixed cost to do so. There is some probability that forecasting is prohibitively expensive, i.e., forecasting does not add sufficient profits to the integrated system to cover the retailer's cost. The supplier is unable confirm whether the retailer can forecast, whether he has forecasted, or what he has learned. She must therefore design a menu of contracts that balances profits earned from a non-forecasting retailer with profits from a forecasting retailer.

We consider buy back and quantity flexibility (QF) contracts. Both are partial returns policies. Under a buy back contract, the retailer may return any amount of unsold stock for a partial refund of the purchase price. Under a QF contract, the retailer may return only a limited amount of unsold stock but receives a full refund of the purchase price. Thus, a buy back contract is a price-based mechanism with a higher return rate representing greater flexibility while a QF contract is a quantity-based mechanism with a higher quantity limit representing greater flexibility. Pasternack (1985) shows that a buy back policy can coordinate the system (i.e., allow a decentralized supply chain to earn the profit of a centralized one) and arbitrarily divide profits between the players. Tsay (1999) establishes similar results for QF contracts. Little has been done on whether one type of contract is preferable to the other. We examine which has a relative advantage in inducing a retailer to forecast.

We first show that one cannot use coordinating buy back contracts to distinguish between forecasting and non-forecasting retailers. If one desires a menu of contracts such that a retailer who is capable of forecasting accepts one contract while a non-forecasting retailer accepts a distinct contract, then one cannot use coordinating buy back contracts. To assure separation between the retailers using buy backs, one must sacrifice supply-chain efficiency for at least one type of retailer. With QF contracts, on the other hand, the supplier can induce separation with coordinating contracts, but her profit may be maximized by offering at least one contract that does not coordinate the supply chain. QF contracts would consequently appear to be better suited for providing incentives to forecast. It is therefore surprising that there are many instances in which the supplier is better off using buy backs. The cost of forecasting is key in determining which contract form is better. When forecasting is relatively cheap, buy backs are preferred. When forecasting is expensive, QF contracts preform better.

The role of the information on contracting has been an important topic in economics and marketing. Recently, there has been some work in the supply-chain contracting literature. (See Cachon, 2001, and Chen, 2001a, for reviews.) Donohue (2000) examines a setting with two production modes, an inexpensive one with a long lead time and an expensive mode with a short lead time. A signal regarding demand (possibly imperfect) is received when only the expensive mode is available. She shows that the basic results of Pasternack (1985) go through: buy backs can coordinate the system and arbitrarily divide profits. Tsay (1999) also considers the possibility of an imperfect demand signal. Our setting differs from these models in that they suppose the demand signal happens automatically and is seen by all; we suppose that the retailer must choose to forecast and only he observes the outcome.

In Cachon and Lariviere (2001a), a manufacturer is privately informed about demand and must contract with a supplier to provide capacity for a critical component. They show that asymmetric information may lead a manufacturer expecting a large market to offer terms that she would never offer under full information. In particular, she may commit to a minimum purchase level to demonstrate that her claim of a big market is credible. Here, the uniformed party offers the contract. Therefore, we concentrate on contracts that "screen" informed parties as opposed to contracts that "signal" to uninformed parties.

Several recent papers have examined screening in a supply-chain setting. Corbett and de Groote (2000), Corbett (2001), and Ha (2001) consider screening based on asymmetric cost information (e.g., buyer holding cost) but assume that the demand distribution is common knowledge. We have asymmetric information on the cost of forecasting but this leads to asymmetric information about the demand distribution (assuming forecasting takes place). Chen (2001b) examines an incentive scheme that induces sales personnel to reveal the potential demand of their assigned territory. In Porteus and Whang (1999), a buyer knows the size of his market, and a supplier posts a menu of contracts that induces the buyer to reveal this information. Because a buyer knows his market size, offering contracts with different absolute minimum purchase requirements is an effective screening device. In our model, not all forecasts are good news; a forecasting retailer may learn that the market is, in fact small. An absolute minimum purchase quantity is consequently not an attractive contract requirement. Rather, we consider QF contracts that feature a proportional minimum purchase.

Below, we first present the model and the contracts. §3 then develops a model of forecasting. §4 examines the performance of a decentralized supply chain. Finally, §5 discusses the results and possible generalizations. Unless otherwise stated, proofs are in the Appendix.

2 Model basics and contracts

We introduce the basics of the model in the context of an integrated supply chain and then turn to a decentralized structure and contracts.

2.1 A centralized system

The supply chain sells one product for which demand is random. There is a single selling season and a single opportunity to produce the good before demand is realized. The marginal cost of producing the good is c, and it sells at a fixed retail price r. Both salvage and overage costs are set to zero. An integrated supply chain thus face a newsvendor problem.

Let $\Phi_0(\xi)$ denote the initial estimate of the demand distribution and $\phi_0(\xi)$ the corresponding density. $\bar{\Phi}_0(\xi) = 1 - \Phi_0(\xi)$. If no forecasting takes place, the profit of the integrated system given a stocking level y is:

$$\pi_0(y) = -cy + r\int_0^y \xi\phi_0(\xi)\, d\xi + ry\bar{\Phi}_0(y), \tag{1}$$

and the optimal stocking level y_0^I satisfies

$$\Phi_0\left(y_0^I\right) = \frac{r-c}{r}. \tag{2}$$

If forecasting takes place, a signal σ (possibly vector valued) is observed. The demand distribution is updated to $\Phi_1(\xi|\sigma)$ with density $\phi_1(\xi|\sigma)$ and $\bar{\Phi}_1(\xi|\sigma) = 1 - \Phi_1(\xi|\sigma)$. Integrated system profit (gross of forecasting costs) and the optimal stocking level y_σ^I are defined analogously to (1) and (2), respectively. Let π_0^I denote optimal expected profit without forecasting, π_σ^I optimal expected gross profit given an observed demand signal σ, and π_1^I optimal expected gross profit given that forecasting will take place. $\pi_1^I = E\left[\pi_\sigma^I\right]$, where the expectation is taken over possible signals σ. For now, we assume that $\pi_1^I > \pi_0^I$. In §3, we present conditions such that this holds.

Forecasting costs κ. κ equals κ_F with probability ρ and κ_N with probability $1-\rho$ for $0 < \rho < 1$. Its value is observed before forecasting and ordering take place. We assume:

$$0 \leq \kappa_F < \pi_1^I - \pi_0^I \leq \kappa_N.$$

Thus if $\kappa = \kappa_F$, the system should forecast since it expects to cover its cost. If $\kappa = \kappa_N$, the supply chain should not forecast as it will not (in expectation) recoup its cost.

2.2 A decentralized system

In a decentralized supply chain, an upstream supplier must sell through a downstream retailer. The supplier is the leader in the channel and gets to set the terms of trade. The supplier incurs the cost of production while the retailer incurs the cost of forecasting and collects revenue from the market. The parties agree on the initial demand distribution and on the probability ρ that the retailer can forecast, but only the retailer observes the realized forecasting cost κ. If $\kappa = \kappa_F$, the supplier cannot observe if the retailer forecasts. If he does forecast, the supplier cannot observe the realized value of σ.

To achieve channel coordination, we first need to assure that forecasting takes place only when appropriate. Next we need to assure that the correct stocking level – y_0^I or y_σ^I – is chosen. How the total stocking level for the supply chain is set is therefore critical. This, as well monetary transfers between the parties, depends on the terms governing transactions. We next discuss possible contracts beginning with buy backs. To simplify notation we drop subscripts from distributions and related functions and write, for example, $\Phi(\xi)$.

2.2.1 Buy back contracts

A buy back contract consists of two parameters, a wholesale price w and a buy back rate b. $r > w \geq c$ and $w > b \geq 0$. The supplier posts the terms, and the retailer determines the order quantity y, paying wy to the supplier. The retailer keeps all revenue from selling the product. If realized demand is $D < y$, the supplier pays $b(y - D)$ to the retailer.[1] All aspects of the contract are assumed enforceable; the retailer must pay for his order, and the supplier must buy back excess stock. For more on this point, see Cachon (2001).

Given this contract, the retailer faces a newsvendor problem with acquisition cost w and salvage value b. His profit may be written as:

$$\Pi_R(y|w,b) = -wy + r\int_0^y \xi\phi(\xi)\,d\xi + ry\bar{\Phi}(y) + b\int_0^y (y-\xi)\,\phi(\xi)\,d\xi,$$

and his optimal order y^* satisfies $\Phi(y^*) = (r-w)/(r-b)$. Let $\Pi_R^*(w,b) = \Pi_R(y^*|w,b)$. The supplier's expected profit is then

$$\Pi_S^*(w,b) = (w-c)\,y^* - b\int_0^{y^*} (y^*-\xi)\,\phi(\xi)\,d\xi.$$

The following lemma presents some basic facts about buy back contracts. It requires the following definition (Shaked and Shanthikumar, 1994): For two random variables X_0 and X_1, we say that X_1 is smaller than X_0 in the convex order and write $X_1 \leq_{cx} X_0$ if $E[\psi(X_1)] \leq E[\psi(X_0)]$ for all convex functions ψ.

[1] Whether the supplier literally "buys back" unsold stock or merely supplements the retailer's salvage opportunities is immaterial. For alternative ways of implementing buy backs, see Lariviere (1999) and Cachon and Lariviere (2001b).

Lemma 1 *Suppose there are two independent markets with demand random variable in market i being X_i $i = 0, 1$. The supplier offers a buy back contract $\{w, b\}$ to a retailer in each market, and retailer i orders optimally given this contract and X_i. (The retailers do not compete for customers.) Let y^i be retailer i's optimal order given $\{w, b\}$ and Π_R^i [Π_S^i] be the resulting retailer [supplier] profit.*

1. If $X_1 \leq_{cx} X_0$, then $\Pi_R^1 \geq \Pi_R^0$.
2. If $X_1 = \theta X_0$ for $\theta > 0$, $y^1 = \theta y^0$, $\Pi_R^1 = \theta \Pi_R^0$, $\Pi_S^1 = \theta \Pi_S^0$.

The convex ordering is a variability ordering. It implies that the smaller random variable has a smaller variance and smaller coefficient of variation. The first part of the lemma is a generalization of Gerchak and Mossman (1992) and shows that the retailer is better off with less uncertainty. The second part gives a useful relationship between markets that differ only on a scale parameter. By considering $\{w, b\} = \{c, 0\}$, one can generalize the results to a centralized supply chain.

The preceding lemma may be interpreted as allowing a retailer to choose between markets. The next lemma considers a retailer choosing between contracts.

Lemma 2 *Suppose that the retailer must choose between two contracts $\{w_1, b_1\}$ and $\{w_2, b_2\}$ such that $w_1 > w_2$ and $\frac{r-w_1}{r-b_1} \leq \frac{r-w_2}{r-b_2}$. The retailer prefers $\{w_2, b_2\}$.*

The lemma implies that a retailer is only interested in a contract with a high wholesale price if it offers significant more flexibility, i.e., a return rate generous enough to move the retailer to a higher critical fractile. If contract 1 has a higher price but results in a higher critical fractile, then whether the retailer prefers contract 1 or 2 depends on the demand distribution he faces. This observation will be useful in designing a menu of contracts.

We now characterize buy back contracts that coordinate the supply chain. The proof of the following can be found in Pasternack (1985).

Theorem 3 *Suppose the supplier offers $\{w_\varepsilon, b_\varepsilon\}$ for $0 \leq \varepsilon < r - c$ where*

$$w_\varepsilon = c + \varepsilon \quad \text{and} \quad b_\varepsilon = \varepsilon \frac{r}{r-c}.$$

Then the retailer orders the integrated channel quantity, i.e., $y^* = y^I$. The retailer profit is $\Pi_R^*(w_\varepsilon, b_\varepsilon) = \left(1 - \frac{\varepsilon}{r-c}\right) \pi^I$. The supplier profit is $\Pi_S^*(w_\varepsilon, b_\varepsilon) = \frac{\varepsilon}{r-c} \pi^I$.

Theorem 3 is among the most commonly cited results in the supply-chain contracting literature. Three features of coordinating buy back contracts are noteworthy. First, a party's profit is increasing in its responsibility for unsold stock. While the retailer prefers a higher back rate for a fixed wholesale price, he prefers the coordinating buy back with the lowest possible buy back rate ($b_0 = 0$) since it gives him the lowest wholesale price. Second, there exist a continuum coordinating contracts that only differ with regard to how the supply chain profit is divided. Consequently, any split of profit can be achieved without any side payments. Finally, the contract $\{w_\varepsilon, b_\varepsilon\}$ is independent of the demand distribution Φ. The same contract will both coordinate the system and yield the same profit split for any Φ.

2.2.2 Quantity flexibility contracts

Like buy backs, a quantity flexibility contract $\{w, d\}$ is a two parameter contract. w again represents a wholesale price; d is a return limit.[2] $0 \le d < 1$. The supplier first posts the terms, and the retailer then sets the stocking level y and pays the supplier wy. The retailer keeps all revenue from selling the product. If realized demand D is $y > D \ge y(1-d)$, the retailer returns $y - D \le yd$ units to the supplier for a full refund of the wholesale price (i.e., he receives $w(y-D)$). If demand is less than $y(1-d)$, the retailer may return yd units and receives wyd. The retailer salvages the remaining $y(1-d) - D$ units on his own. The return limit d thus measures the flexibility that the supplier offers with a higher value representing greater flexibility. The supplier takes responsibility for the first yd unsold units, and the retailer takes responsibility for the next $y(1-d)$.

Under a QF contract, the retailer's profit can be written as follows:

$$\Pi_R(y|w,d) = (r-w)\int_0^y \xi\phi(\xi)\,d\xi + (r-w)\,y\bar{\Phi}(y) - w\int_0^{y(1-d)} (y(1-d)-\xi)\,\phi(\xi)\,d\xi. \quad (3)$$

$\Pi_R(y|w,d)$ is concave in y and the optimal retailer order y^* must satisfy:

$$(r-w)\,\bar{\Phi}(y^*) = w(1-d)\,\Phi(y^*(1-d)). \quad (4)$$

The left-hand side may be interpreted as the retailer's expected marginal benefit of increasing his order while the right-hand side represents his expected marginal cost. Clearly, if $d = 0$, (3) reduces to a standard newsvendor problem, and (4) becomes the usual critical fractile solution. For $d > 0$, the problem is somewhat harder since there is generally no explicit expression of y^* even if Φ is easily invertible. However, the concavity of $\Pi_R(y|w,d)$ allows one to determine y^* numerically fairly simply. Additionally, implicit differentiation shows

---

[2] Tsay (1999) and Tsay and Lovejoy (1999) define QF contracts as having three parameters. In our one-period setting, the third parameter is superfluous.

that y^* is increasing in d. Let $\Pi_R^*(w,d) = \Pi_R(y^*|w,d)$. The corresponding supplier profit is

$$\Pi_S^*(w,b) = -cy^* + wy^*(1-d)\Phi(y^*(1-d)) + w\int_{y^*(1-d)}^{y^*} \xi\phi(\xi)\,d\xi + wy^*\bar{\Phi}(y).$$

The following are the QF counterparts to Lemmas 1 and 2.

Lemma 4 *Suppose there are two independent markets with demand random variable in market i being X_i $i = 0, 1$. The supplier offers a QF contract $\{w, d\}$ to a retailer in each market, and retailer i orders optimally given this contract and X_i. (The retailers do not compete for customers.) Let y^i be retailer i's optimal order given $\{w, d\}$ and Π_R^i [Π_S^i] be the resulting retailer [supplier] profit.*

1. If $X_1 \leq_{cx} X_0$, then $\Pi_R^1 \geq \Pi_R^0$.
2. If $X_1 = \theta X_0$ for $\theta > 0$, $y^1 = \theta y^0$, $\Pi_R^1 = \theta \Pi_R^0$, $\Pi_S^1 = \theta \Pi_S^0$.

Lemma 5 *Suppose that the retailer must choose between two contracts $\{w_1, d_1\}$ and $\{w_2, d_2\}$ such that $w_1 > w_2$ and $y^1 \leq y^2$ where y^i is the optimal order under contract $i = 1, 2$. The retailer prefers $\{w_2, d_2\}$.*

We again have that the retailer is better off in a less variable market and is only interested in a contract with a higher wholesale price if it results in a higher stocking level. One caveat to the latter result that is different from the buy back case is that not all retailers will evaluate two contracts the same way. With buy backs, all retailers will pick the same critical fractile regardless of their demand distribution. With QF contracts, one retailer given two contracts may choose $y^1 > y^2$ while another retailer facing a different demand distribution may choose $y^1 < y^2$. An example illustrates this possibility. Suppose demand in market j for $j = 1, 2$ follows a power function distribution with parameter k_j. $\Phi_j(\xi) = \xi^{k_j}$ for $0 \leq \xi \leq 1$ and $k_j > 0$. Assume $k_1 = 1/2$ and $k_2 = 2$. The retail price is $r = 10$, $\{w_A, d_A\} = \{5, 0.1\}$, and $\{w_B, d_B\} = \{9, 0.57\}$. Let y_j^i denote the optimal order for market $j = 1, 2$ under contract $i = A, B$.

$$y_j^i = \left(\frac{r - w_i}{r - w_i + w_i(1-d)^{k_j+1}}\right)^{1/k_j}$$

We have:

	y_1^i	$\Phi_1(y_1^i)$	y_2^i	$\Phi_2(y_2^i)$
$\{w_A, d_A\}$	0.290	0.539	0.761	0.578
$\{w_B, d_B\}$	0.080	0.283	0.763	0.583

Thus, while moving from contract A to contract B results in a modest increase in the service level in market 2, it produces a dramatic drop in the service level in the first market.

We now consider coordinating QF contracts. The following is from Tsay (1999).

Theorem 6 *If the supplier offers* $\{w_d, d\} = \left\{\frac{c}{\frac{c}{r}+(1-d)\Phi((1-d)y^I)}, d\right\}$ *for* $0 \leq d < 1$, *the retailer orders the integrated channel quantity, i.e.,* $y^* = y^I$. *The retailer profit is decreasing in* d *with* $\Pi_R^*(w_0, 0) = \pi^I$ *and* $\lim_{d\to 1} \Pi_R^*(w_d, d) = 0$. *The supplier's profit is increasing in* d.

Of the three properties of coordinating buy backs identified above, two carry over to coordinating QF contracts. A continuum coordinating contracts exists, and the contracts only differ in how they split the supply chain profit. A party's profit is increasing in the responsibility it takes for excess inventory. The property that does not carry over is that coordinating QF contracts depend on the demand distribution. One cannot simply write down a single contract that works for all markets. As a consequence, to understand whether one is better off using buy backs or QF contracts, one must first consider how QF contracts depend on the demand distribution. We now examine this question.

2.2.3 The dependence of QF contracts on the demand distribution

We begin with a definition. Given random variables X_0 and X_1 with respective distribution Φ_0 and Φ_1 and respective inverses Φ_0^{-1} and Φ_1^{-1}, let $G(\xi) = \Phi_0^{-1}(\Phi_1(\xi))/\xi$. We say that X_1 is smaller than X_0 in the *star order* and write $X_1 \leq_* X_0$ if $G(\xi)$ is increasing (Shaked and Shanthikumar, 1994). Many common distributional families can be ordered according to the star ordering. Consider the Weibull distribution with density

$$\phi(\xi|\theta, k) = \theta k \xi^{k-1} e^{-\theta\xi^k} \text{ for } \theta > 0, k > 0, \xi > 0 \tag{5}$$

or the gamma distribution with density

$$\eta(\xi|\theta, k) = \theta^k \xi^{k-1} e^{-\theta\xi}/\Gamma(k) \text{ for } \theta > 0, k > 0, \xi > 0. \tag{6}$$

For either family, the parameter k is referred to as the shape parameter. Barlow and Proschan (1975) show that if X_k is a Weibull or gamma random variable with shape parameter k, then $X_{k_1} \leq_* X_{k_0}$ if $k_1 \geq k_0$. Lariviere and Porteus (2001) show that if $X_i = \alpha_i + \beta_i X$ for some random variable X and parameters $\alpha_i \geq 0$ and $\beta_i > 0$, then $X_1 \leq_* X_0$ if $\alpha_1/\beta_1 \geq \alpha_0/\beta_0$. Consequently, the normal and uniform families are easily ordered by the star order.

We now present alternative characterizations and some useful properties of the star order.

Theorem 7 *Suppose that* X_0 *and* X_1 *are non-negative.* X_1 *is smaller than* X_0 *in the star order if and only if*

1. $\frac{\Phi_0^{-1}(\beta)}{\Phi_0^{-1}(\alpha)} \geq \frac{\Phi_1^{-1}(\beta)}{\Phi_1^{-1}(\alpha)}$ for $0 < \alpha \leq \beta < 1$.

2. $\Phi_0\left(\Phi_0^{-1}(\alpha)\lambda\right) \geq [\leq] \Phi_1\left(\Phi_1^{-1}(\alpha)\lambda\right)$ for $\alpha \in (0,1)$ and $\lambda < [>] 1$.
3. Assuming X_0 and X_1 have respective densities ϕ_0 and ϕ_1, $\Phi_0^{-1}(\alpha)\phi_0\left(\Phi_0^{-1}(\alpha)\right) \leq \Phi_1^{-1}(\alpha)\phi_1\left(\Phi_1^{-1}(\alpha)\right)$ for $\alpha \in (0,1)$.

Theorem 8 *Suppose that $X_1 \leq_* X_0$.*

1. Let $\tilde{X}_i = X_i/E[X_i]$. Then $\tilde{X}_1 \leq_{cx} \tilde{X}_0$.
2. $\theta_1 X_1 \leq_* \theta_0 X_0$ for $\theta_0, \theta_1 > 0$.
3. For $\alpha \in (0,1)$, Φ_0^{-1} and Φ_1^{-1} can cross at most once. If they do cross at, say, α^*, then $\Phi_0^{-1}(\alpha) < [>] \Phi_1^{-1}(\alpha)$ for $\alpha < [>] \alpha^*$.

The star order is a variability ordering. The larger random variable X_0 limits how fast Φ_1^{-1} may increase as one moves from α to β while Theorem 8 explicitly links the star order to the more familiar convex ordering. The latter also implies that X_1 must have a smaller coefficient of variation than X_0 and (in light of Lemma 1) that a newsvendor makes a higher profit per unit of mean demand in a smaller market. What distinguishes the star order from other variability orders is that it implies little about the random variables' respective magnitudes. For example, the convex ordering implies that the random variables have the same mean. If X_1 is smaller than X_0 in the dispersive ordering[3] (subject to some mild restrictions), it must be that $\Phi_1(\xi) \leq \Phi_0(\xi)$ for all ξ (Shaked and Shanthikumar, 1994). The star order imposes no such restrictions. The star order is useful in a newsvendor setting because it compares distributions at fractiles as opposed to at quantities, and a newsvendor's decisions are dictated by these fractiles. Also, Theorem 7 part 2 allows us to say something about rescaling of fractiles. It is crucial to proving the following.

Theorem 9 *Suppose retailers 0 and 1 are in independent markets and face stochastic demands X_0 and X_1, respectively. The supplier offers both a QF contract $\{w, d\}$ with $d > 0$. Φ_j is the distribution of X_j and y_j the optimal order for retailer $j = 0, 1$. If $X_1 \leq_* X_0$, then:*

1. Retailer 1 chooses a higher service level, i.e., $\Phi_1(y_1) \geq \Phi_0(y_0)$.
2. Retailer 1 is more likely to exploit the flexibility the supplier offers, i.e., $\Phi_1(y_1(1-d)) \leq \Phi_0(y_0(1-d))$.

The theorem suggests that if a retailer faces a smaller market (in the sense of the star ordering), he is better able to exploit a QF contract. In particular, he finds committing to

[3] X_1 is smaller than X_0 in the dispersive order ($X_1 \leq_{disp} X_0$) if $\Phi_0^{-1}(\beta) - \Phi_0^{-1}(\alpha) \geq \Phi_1^{-1}(\beta) - \Phi_1^{-1}(\alpha)$ for $0 < \alpha \leq \beta < 1$.

a minimum purchase less onerous because his demand is more likely to be greater than the level to which he has committed. Consequently, he can afford to increase his stocking level and serve a higher fraction of demand.

Theorem 10 *Suppose retailers* 0 *and* 1 *are in independent markets and face stochastic demands* X_0 *and* X_1*, respectively. The supplier offers coordinating contracts with return limit* $d > 0$ *in both markets. Let* w_d^j *be the corresponding wholesale price in market* $j = 0, 1$.

1. If $X_1 \leq_* X_0$, $w_d^1 \geq w_d^0$.
2. If $X_1 = \theta X_0$ for $\theta > 0$, $w_d^1 = w_d^0$.

Under a QF contract with $d > 0$, the retailer in the smaller market sets a higher service level for a fixed contract. But coordination in a newsvendor setting is all about getting the retailer to pick choose the correct service level. A higher wholesale price is consequently needed in the smaller market to rein in the retailer. The second part of the theorem demonstrates that while a coordinating QF contract depends on the demand distribution, it does not necessarily depend on all aspects of the distribution. In particular, a scale parameter is irrelevant in determining the coordinating wholesale price.

Theorem 10 says nothing about how the retailer's share of supply-chain profit for a given d varies between the markets. Making such a general statement is difficult since one does not necessarily have a simple expression for retailer profit under a coordinating QF contract. However, the case of power function demand is simple. Let X_k have a power function distribution with parameter $k > 0$. It is easy to verify that $X_{k_1} \leq_* X_{k_0}$ for $k_1 > k_0$. The integrated integrated system profit is $\pi_k^I = \mu_k y_k^I (r - c)$ where $\mu_k = E[X_k]$ and $y_k^I = (1 - c/r)^{1/k}$. The coordinating wholesale price is $w_d(k) = r\left(1 + (1-d)^{k+1} \frac{(r-c)}{c}\right)^{-1}$. If the supplier offers $\{w_d(k), d\}$, the resulting retailer profit is

$$\Pi_R^k = \frac{r}{c(1-d)^{-k-1} + (r-c)} \pi_k^I.$$

Thus, the retailer's share of system profit is decreasing in k for a fixed d.

3 A model of forecasting

We now present how the supply chain forecasts. We assume the supply chain knows the family of the demand distribution and but not some parameter of the distribution. For example, it knows that demand is exponentially distributed but not the rate of the exponential. The firm, however, has a prior on the missing parameter. If the supply chain were able to see a demand draw, it could update its beliefs using Bayes formula in the manner of Scarf (1959).

Consequently, we model the forecasting signal σ as n "faux" demand realizations (for $n = 1, 2, \ldots$). Letting ξ_i denote the i^{th} draw, $\sigma = \{\xi_1, \ldots, \xi_n\}$. The draws are real in the sense that they come from the true demand distribution. They are fake in the sense that no actual transactions take place; the supply chain need not hold stock to serve these realizations, and no cash changes hands. The n draws are assumed independent of each other and of the ultimate real draw from the market place. While independent of the eventual demand realization of demand, the faux draws provide useful information since they allow the firm to refine its estimate of the missing parameter. We interpret the number of draws as a measure of the effectiveness of the forecasting technology with a higher value corresponding to a better technology. Note that n is exogenously specified. The decision is whether or not to forecast, not how precise a forecast to obtain.

Let $\phi(\xi|\theta)$ denote the density of the true demand distribution given θ. The true value of θ is unknown. Let $\eta_0(\theta)$ denote the prior density. For simplicity, assume η_0 is taken from a conjugate family of ϕ (DeGroot, 1970). Averaging over possible values of θ, we have:

$$\phi_0(\xi) = \int_{-\infty}^{\infty} \phi(\xi|\theta)\,\eta_0(\theta)\,d\theta.$$

We call the $\phi_0(\xi)$ the predictive density given the prior. If no forecasting takes place, it is the supply chain's estimate of the demand distribution. Now suppose that forecasting does take place. Because we are working with a conjugate family, the posterior distribution of θ given σ, $\eta_1(\theta|\sigma)$, is from the same family as the initial prior. Additionally, the updated predictive density $\phi_1(\xi) = \int_{-\infty}^{\infty} \phi(\xi|\theta)\,\eta_1(\theta|\sigma)\,d\theta$ is from the same family as $\phi_0(\xi)$.

Given $\phi_1(\xi)$, it is straightforward to evaluate the integrated supply chain's optimal order y_σ^I and the resulting system profit π_σ^I. For a decentralized supply chain governed by a buy back or QF contract, it is also straightforward to determine either the supplier's profit given σ, Π_σ^S, or the corresponding retailer's profit, Π_σ^R. However, to determine whether the integrated supply chain should forecast, we require $\pi_1^I = E\left[\pi_\sigma^I\right]$. Similarly, in a decentralized supply chain, we need $E\left[\Pi_\sigma^S\right]$ to address what contract the supplier should offer and $E\left[\Pi_\sigma^R\right]$ to address whether the retailer should forecast. Unfortunately, y_σ^I and π_σ^I are generally complex functions of the observed market signal (particularly when $n > 1$), so ascertaining π_1^I is not simple. Calculations for the decentralized case are even more complex. We must make some simplifying assumptions to move forward; we turn to the state space reduction introduced by Scarf (1960) and extended by Azoury (1985).

We henceforth assume that $\phi(\xi|\theta)$ is a Weibull density as defined in (5) with known shape parameter k and unknown scale parameter θ. The gamma is a conjugate distribution of the Weibull (Azoury, 1985), so we assume that $\eta_0(\theta)$ is gamma density as defined in (6) with scale parameter S_0 and shape parameter a_0. Assume $a_0 k > 1$. Given market signal σ, the posterior is a gamma distribution with parameters S_1 and a_1 where:

$$S_1 = S_0 + \sum_{j=1}^{n} \xi_j^k \quad \text{and} \quad a_1 = a_0 + n.$$

Only the scale parameter of the posterior S_1 depends on the observed signal. The shape parameter a_1 depends only on the fact that forecasting has occurred. The coefficient of variation of a gamma distribution with shape parameter a is $1/\sqrt{a}$. Hence, "a" stands for accuracy; the shape parameter represents the precision of information with a higher value corresponding to lower uncertainty in the supply chain's estimate of the missing parameter. Forecasting reduces the uncertainty of the estimate regardless of the actual signal observed.

For a Weibull distribution with a gamma prior, the predictive distribution is a Burr Type XII distribution (hereafter, simply a Burr distribution; see Burr, 1942). Given the Weibull parameter k and the gamma parameters S and a, $\phi_i(\xi) = \phi(\xi|S_i, a_i)$ for $i = 0, 1$ where:

$$\phi(\xi|S, a) = \frac{akS^a \xi^{k-1}}{\left(1 + \xi^k\right)^{a+1}}. \tag{7}$$

Lemma 11 *Let X_i be a Burr distribution with parameters (S_i, a_i) for $i = 0, 1$.*

1. If $a_i k > j$, then the j^{th} moment of X_i equals $a_i S_i^{j/k} \frac{\Gamma(a_i - j/k)\Gamma(1+j/k)}{\Gamma(a_i+1)}$.
2. If $a_1 = a_0$, then $X_1 = S_1^{1/k} S_0^{-1/k} X_0$.
3. If $a_1 \geq a_0$, $X_1 \leq_* X_0$.

Thus $\phi_0(\xi)$ and $\phi_1(\xi)$ are related by the star order. Let $\mu(S, a)$ denote the mean of a Burr distribution with parameters (S, a). By our earlier assumption $a_0 k > 1$, the mean of the predictive distribution before forecasting, $\mu(S_0, a_0)$, is in fact well defined. Let $\tilde{X}_a$ be a Burr random variable with parameters $\left(\mu(1, a)^{-k}, a\right)$. (Note that the mean of $\tilde{X}_a$ is one.) By part 2 of the lemma, if X has a Burr distribution with parameters (S, a), then $X = \mu(S, a)\tilde{X}_a$. The next theorem links the profit under X and $\tilde{X}_a$.

Theorem 12 *Let X and $\tilde{X}_a$ have Burr distributions with respective parameters (S,a) and $\left(\mu(1,a)^{-k}, a\right)$. Let π^I [$\tilde{\pi}_a^I$] denote the optimal integrated supply-chain profit under demand X [$\tilde{X}_a$]. Similarly, for a given buy back contract, let Π_R^B [$\tilde{\Pi}_R^B$] and Π_S^B [$\tilde{\Pi}_S^B$] be, respectively, the retailer and supplier profit in a decentralized supply chain facing demand X [$\tilde{X}_a$]. Let Π_R^Q, $\tilde{\Pi}_R^Q$, Π_S^Q, and $\tilde{\Pi}_S^B$ be the corresponding quantities for a given QF contract.*

1. $\pi^I = \mu(S,a)\,\tilde{\pi}_a^I$. $\tilde{\pi}_a^I$ is increasing in a.
2. $\Pi_R^B = \mu(S,a)\,\tilde{\Pi}_R^B$ and $\Pi_S^B = \mu(S,a)\,\tilde{\Pi}_S^B$. $\tilde{\Pi}_R^B$ is increasing in a.
3. $\Pi_R^Q = \mu(S,a)\,\tilde{\Pi}_R^Q$ and $\Pi_S^Q = \mu(S,a)\,\tilde{\Pi}_S^Q$. $\tilde{\Pi}_R^Q$ is increasing in a.

We refer to $\tilde{\pi}_a^I$ as the normalized return. It represents the integrated system's profit per unit of mean demand. Writing profits in this manner clarifies how forecasting improves system performance. Given a realized market signal σ and resulting distribution parameters (S_1, a_1), the resulting mean $\mu(S_1,a_1)$ may be more or less than the initial estimate of the mean $\mu(S_0,a_0)$. However, the post-forecasting normalized return $\tilde{\pi}_{a_1}^I$ is certain to be higher. Thus, even if the realized forecast suggests that the market is smaller than initially thought, total expected profit may be higher if the normalized return increases sufficiently.

The normalized returns can be determined up front, making calculating expected profits relatively simple. In particular, $\pi_\sigma^I = \mu(S_1,a_1)\,\tilde{\pi}_{a_1}^I$, so:

$$\pi_1^I = E\left[\mu(S_1,a_1)\,\tilde{\pi}_{a_1}^I\right] = E\left[\mu(S_1,a_1)\right]\tilde{\pi}_{a_1}^I.$$

To determine whether forecasting is profitable, we need only examine $E\left[\mu(S_1,a_1)\right]$.

Lemma 13 $E\left[\mu(S_1,a_1)\right] = \mu(S_0,a_0)$.

At an intuitive level, the lemma is appealing. It says that one does not expect the size of the market to change if one forecasts. Of course, one expects the variance to fall. Using the star ordering, one can show that this is fact the case. We had earlier assumed that forecasting increased the profitability of the system; we now have that this is a consequence of our forecasting structure and conclude that the integrated supply chain should forecast if the cost of forecasting κ satisfies the following:

$$\frac{\kappa}{\mu(S_0,a_0)} \leq \tilde{\pi}_{a_1}^I - \tilde{\pi}_{a_0}^I.$$

4 Performance of a decentralized supply chain

We now consider a decentralized supply chain that faces the forecasting and stocking problem modeled above. We assume the retailer is privately informed of his cost to forecast κ. The supplier has a prior $0 < \rho < 1$ that $\kappa = \kappa_F$ (i.e., the retailer can afford to forecast). The players agree on the initial estimate of the market as captured by the parameters (S_0, a_0). The sequence of events is:

1. The supplier posts a menu of either buy back or QF contracts.
2. The retailer learns his cost of forecasting κ.
3. The retailer agrees to carry the product (incurring cost τ) if he expects to turn a profit.
4. He commits to a contract and then forecasts (incurring cost κ) if doing so maximizes his profit.
5. The retailer orders stock and pays the supplier as dictated by the contract he has chosen.
6. Market demand is realized.
7. Returns as allowed by the retailer's chosen contract take place.

There are some points to note about the assumed sequence. First, we restrict the supplier to offer only buy back contracts or only QF contracts. She cannot design a menu which offers, say, a buy back to the forecaster and a QF contract to the non-forecaster. This simplifies the analysis and allows us to focus on the relative performance of the two types of contracts. Second, we assume that the retailer incurs an opportunity cost τ when he commits to carry the product. This represents the costs of handling the inventory, making space on shelves etc. as well as the profit foregone by carrying the supplier's product instead of something else. It does not include the cost of forecasting, κ. We assume that $\tau < \pi_0^I$; it is thus efficient for the integrated system to stock the product even if no forecasting takes place. The retailer will only carry the product if he expects at stage 3 a return greater than his opportunity costs. However, because the opportunity cost τ is sunk, he will still offer the product to the market if he subsequently learns from forecasting that the market will be very small. Finally, we have assumed that the retailer must commit to a contract before forecasting. An alternative formulation of the fourth step may be more realistic in some settings:

4′. The retailer forecasts if doing so maximizes his profit and then commits to a contract.

As explained below, our formulation can handle this alternative assumption.

4.1 The supplier's contract design problem

We begin with some notation. We use a superscript R [S] to denote values relevant to the retailer [supplier]. A subscript F [N] denotes values relevant to a retailer who has [not] forecasted. Arguments of function will be contract terms, $\{w, b\}$ in the case of a buy back contract and $\{w, d\}$ in the case of a QF contract. Thus, $\Pi_F^S(w_F, d_F)$ is the supplier's profit when the retailer has forecasted and accepted QF $\{w_F, d_F\}$ and $\Pi_F^R(w_N, b_N)$ is the retailer's profit when he has forecasted and accepted buy back contract $\{w_N, b_N\}$. As in section 3, we use a tilde to denote normalized returns.

The supplier's problem in the case of offering buy backs can now be stated as follows:

$$\max_{w_F, b_F, w_N, b_N} \rho \Pi_F^S(w_F, b_F) + (1-\rho)\Pi_N^S(w_N, b_N) \qquad (\mathbf{P1})$$

Subject to:

$$\Pi_F^R(w_F, b_F) \geq \Pi_F^R(w_N, b_N) \qquad (\text{IC-}F)$$

$$\Pi_N^R(w_N, b_N) \geq \Pi_N^R(w_F, b_F) \qquad (\text{IC-}N)$$

$$\Pi_F^R(w_F, b_F) \geq \tau + \kappa_F \qquad (\text{IR-}F)$$

$$\Pi_N^R(w_N, b_N) \geq \tau \qquad (\text{IR-}N)$$

The formulation for QF contracts would be identical except that $\{w_j, d_j\}$ would replace $\{w_j, b_j\}$ for $j = F, N$. Here we have implicitly relied on the revelation principle, which (loosely) says that nothing is lost by restricting the analysis to a direct mechanism in which the retailer truthfully reveals his type by his contract selection (Salanié, 1997). The supplier thus offers one contract $\{w_F, b_F\}$ meant for the forecasting retailer and another $\{w_N, b_N\}$ meant for the non-forecaster. The *incentive compatibility* constraints (IC) assure that truth-telling is indeed optimal for both types of retailers. Constraint (IC-F) says that the forecaster prefers the contract intended for him while (IC-N) assures the non-forecaster prefers $\{w_N, b_N\}$. The last two constraints are *individual rationality* constraints (IR) and assure both types of retailers are willing to carry the supplier's product.

Following our analysis of the forecasting model, an equivalent formulation is:

$$\max_{w_F, b_F, w_N, b_N} \rho \tilde{\Pi}_F^S(w_F, b_F) + (1-\rho)\tilde{\Pi}_N^S(w_N, b_N) \qquad (\tilde{\mathbf{P}}\mathbf{1})$$

Subject to:

$$\tilde{\Pi}_F^R(w_F, b_F) \geq \tilde{\Pi}_F^R(w_N, b_N) \qquad (\text{IC-}\widetilde{F})$$

$$\tilde{\Pi}_N^R(w_N, b_N) \geq \tilde{\Pi}_N^R(w_F, b_F) \qquad (\text{IC-}\widetilde{N})$$

$$\tilde{\Pi}_F^R(w_F, b_F) \geq \tilde{\tau} + \tilde{\kappa}_F \qquad (\text{IR-}\widetilde{F})$$

$$\tilde{\Pi}_N^R(w_N, b_N) \geq \tilde{\tau} \qquad (\text{IR-}\widetilde{N})$$

where $\tilde{\tau} = \tau / \mu(S_0, a_0)$ and $\tilde{\kappa}_F = \kappa_F / \mu(S_0, a_0)$. We can thus work with the simpler, normalized returns. It is this ability to move from the general formulation **P1** to the normalized $\tilde{\mathbf{P}}\mathbf{1}$, that allows us to handle the alternative sequencing assumption $4'$. Under this assumption, the forecaster's incentive compatibility constraint (IC-F) would become

$$\Pi_F^R(w_F, b_F | S_1, a_1) \geq \Pi_F^R(w_N, b_N | S_1, a_1) \text{ for all } S_1 > 0. \tag{8}$$

where $\Pi_F^R(w_j, b_j | S_1, a_1)$ denotes retailer profit given market signal σ which leads to posterior parameters (S_1, a_1). However, since $\Pi_F^R(w_j, b_j | S_1, a_1) = \mu(S_1, a_1) \tilde{\Pi}_F^R(w_j, b_j)$ for all $S_1 > 0$, (8) reduces to (IC-$\widetilde{F}$).

In many ways, the supplier's problem is a standard contract-design problem as found in the economics literature (Salanié, 1997). Three outcomes are possible. First, the supplier may forego differentiating between the forecaster and the non-forecaster and offer only one contract that is acceptable to both, making the IC constraints moot. Second, she may choose not to do business with the "unfavorable type" (i.e., the non-forecaster) and offer a contract that only the forecaster would accept. That is, she intentionally offers terms that violate (IR-$\widetilde{N}$). Finally, she may choose to offer incentive compatible contracts that satisfy all the constraints, so she may screen a forecaster from a non-forecaster based on the contract the retailer accepts. We focus on this last possibility but will highlight points at which the supplier might opt for the other alternatives.

Below we establish some properties of the supplier's problem and determine which constraints will bind at the supplier's optimal solution. This highlights the trade offs she confronts. We first consider buy backs and then QF contracts. We conclude this section with a numerical example that compares the relative performance of the two contracts.

4.2 Buy back analysis

We begin by considering coordinating buy back contracts as given in Theorem 3.

Theorem 14 *Let $\{w_{\varepsilon_j}, b_{\varepsilon_j}\}$ for $j = F, N$ be coordinating buy back contracts as defined in Theorem 3. If a pair of contracts $\{w_{\varepsilon_F}, b_{\varepsilon_F}\}$ and $\{w_{\varepsilon_N}, b_{\varepsilon_N}\}$ satisfies the constraints of* **P1***, it must be the case $\varepsilon_F = \varepsilon_N$ so $\{w_{\varepsilon_F}, b_{\varepsilon_F}\} = \{w_{\varepsilon_N}, b_{\varepsilon_N}\}$.*

The supplier cannot induce the retailer to reveal whether he has forecasted using coordinating contracts. Gaining separation requires sacrificing efficiency for at least one type. The supplier must offer at least one contract that results in a critical fractile that is distinct from the integrated system critical fractile. Lemma 2 offers some insight into the structure that must be offered. For the retailer's choice of contract to be dependent on the demand distribution he faces, it must be that one contract features a higher wholesale price with a buy back rate sufficiently high that it results in a critical fractile higher than the contract with the lower wholesale price. If it is the case that a forecasting retailer is induced to take the contract with the lower critical fractile, the supplier does in fact induce forecast revelation by restricting returns. To see if this does occur, we need to examine the IC constraints.

Theorem 15 *Suppose retailers* 0 *and* 1 *are in independent markets and face stochastic demands* X_0 *and* X_1*, respectively. The supplier lets both select from a pair of buy back contracts* $\{w_1, b_1\}$ *and* $\{w_2, b_2\}$ *such that* $w_1 > w_2$ *and* $\frac{r-w_1}{r-b_1} > \frac{r-w_2}{r-b_2}$. *Suppose* $X_1 \leq_* X_0$.

1. If retailer 1 is indifferent between the two contracts, retailer 0 prefers contract $\{w_1, b_1\}$.
2. If retailer 0 is indifferent between the two contracts, retailer 1 prefers contract $\{w_2, b_2\}$.

The theorem implies that only one of the IC constraints binds and that a forecasting retailer takes the contract with a lower wholesale price but less generous return terms. The non-forecaster places greater value on a higher return rate and is consequently willing to pay a higher wholesale price in order to assure generous terms on returns. Two points are worthy of note. First, the result holds for a general model of forecasting that leads to the star ordering and not just for model with Weibull demand we have proposed. Second, the result implies that the iso-profit curves of the two types of retailers can cross only once. Hence, it is equivalent to the usual single crossing condition in the agency literature (Salanié, 1997).

Now consider the IR constraints. For the moment, assume that $\tilde{\kappa}_F = 0$. By Lemma 1, $\tilde{\Pi}_F^R(w_N, b_N) \geq \tilde{\Pi}_N^R(w_N, b_N)$ and any contract satisfying the non-forecaster's IR constraint also satisfies the forecaster's. Consequently, the supplier sets $\{w_N, b_N\}$ such that (IR-$\widetilde{N}$) binds and choose $\{w_F, b_F\}$ so that (IC-$\widetilde{F}$) binds. Further $\{w_F, b_F\}$ is a coordinating contract as given in Theorem 3. To see this, consider the latter point first. Given $\{w_N, b_N\}$, the forecaster can always guarantee himself at least $\tilde{\Pi}_F^R(w_N, b_N)$. From the supplier's perspective, the most efficient way to assure the forecaster that much profit is to offer a coordinating contract as this leaves the most residual profit for the supplier. Next, there is no reason to set $\tilde{\Pi}_N^R(w_N, b_N)$ strictly greater than $\tilde{\tau}$. If $\tilde{\Pi}_N^R(w_N, b_N)$ were greater than $\tilde{\tau}$, the supplier could increase w_N and b_N so that $\chi_N = \frac{r-w_N}{r-b_N}$ remains constant. By Lemma 2, both $\tilde{\Pi}_N^R(w_N, b_N)$ and $\tilde{\Pi}_F^R(w_N, b_N)$ must fall. Since a non-forecaster orders the same quantity as before, total system profit with a non-forecaster is unchanged; $\tilde{\Pi}_N^R(w_N, b_N)$ falling implies that the supplier must earn more in the event that the retailer cannot forecast. Further, a lower value of $\tilde{\Pi}_F^R(w_N, b_N)$ increases what the supplier makes in the event the retailer is a forecaster.

The above analysis shows that the profits of the parties are driven by χ_N, the critical fractile the non-forecaster serves. Define $\{w(\chi_N), b(\chi_N)\}$ as:

$$\{w(\chi_N), b(\chi_N)\}\left\{r - \chi_N (r - b(\chi_N)), r - \frac{\tilde{\tau}}{\int_0^{\tilde{\Phi}_0^{-1}(\chi_N)} \xi \tilde{\phi}_0(\xi)\, d\xi}\right\},$$

where $\tilde{\Phi}_0^{-1}$ is the inverse of a Burr distribution with parameters $\left(\mu(1, a_0)^{-k}, a_0\right)$ and $\tilde{\phi}_0$ is the corresponding density. It is straightforward to verify that $\tilde{\Pi}_N^R(w(\chi_N), b(\chi_N)) = \tilde{\tau}$. Let $\tilde{\pi}_{a_0}^I(\chi_N)$ be the normalized profit for the integrated channel if no forecasting takes place and the stocking level is $\tilde{\Phi}_0^{-1}(\chi_N)$. $\tilde{\pi}_{a_0}^I(\chi_N)$ is decreasing in χ_N for $\chi_N > \chi_I = \frac{r-c}{r}$. Let $\tilde{\Pi}_F^R(\chi_N) = \tilde{\Pi}_F^R(w(\chi_N), b(\chi_N))$. By Theorem 15, $\tilde{\Pi}_F^R(\chi_N)$ is decreasing. The supplier's problem now reduces to choosing χ_N to maximize the following:

$$\Pi^S(\chi_N) = \rho\left(\tilde{\pi}_{a_1}^I - \tilde{\Pi}_F^R(\chi_N)\right) + (1-\rho)\left(\tilde{\pi}_{a_0}^I(\chi_N) - \tilde{\tau}\right).$$

The corresponding first order condition yields:

$$(1-\rho)\frac{d\tilde{\pi}_{a_0}^I(\chi_N)}{d\chi_N} = \rho\frac{d\tilde{\Pi}_F^R(\chi_N)}{d\chi_N}. \tag{9}$$

Let χ_N^* denote the solution to (9).

The trade off the supplier faces is now clear. Boosting χ_N lowers the excess rents she must pay a forecasting retailer. By offering a coordinating contract, she captures every penny she takes away from the forecaster. Those pennies do not come for free. The non-forecaster's profit is fixed $\tilde{\tau}$, and setting $\chi_N^* > \chi_I$ lowers the supplier's pay off if the retailer turns out to be a non-forecaster. The extent to which the supplier is willing to distort the non-forecaster stocking decision depends in part on the probability the retailer can forecast. If ρ is low, one would expect little distortion. If ρ is near one, the supplier puts little weight on the drop in her profit with a non-forecaster and instead focuses on reducing the forecaster's profit.

Now consider $\tilde{\kappa}_F > 0$. Clearly if $\tilde{\Pi}_F^R(\chi_N^*) > \tilde{\tau} + \tilde{\kappa}_F$, the supplier again chooses χ_N^*. If instead $\tilde{\Pi}_F^R(\chi_I) \geq \tilde{\tau} + \tilde{\kappa}_F > \tilde{\Pi}_F^R(\chi_N^*)$, much of the above analysis goes through; the supplier keeps increasing χ_N to reduce the forecaster's excess rents but now stops short of χ_N^* since this would violate the forecaster's IR constraint. Let $\hat{\chi}_N$ solve $\tilde{\Pi}_F^R(\hat{\chi}_N) = \tilde{\tau} + \tilde{\kappa}_F$, we have that for $\tilde{\kappa}_F \leq \tilde{\Pi}_F^R(\chi_I) - \tilde{\tau}$, the supplier choose $\max\{\chi_N^*, \hat{\chi}_N\}$.

If $\tilde{\kappa}_F > \tilde{\Pi}_F^R(\chi_I) - \tilde{\tau}$, the problem changes, and the forecaster's IR constraint plays a bigger role. Suppose the supplier only offered a single coordinating contract, $\{w_\varepsilon, b_\varepsilon\}$. He must choose ε so that the forecaster just breaks even.[4] Now the non-forecaster receives a windfall from the supplier's attempts to accommodate the forecaster. The supplier could reduce the non-forecaster's excess rents by distorting the forecaster's critical fractile χ_F. By Theorem 15, she must set $\chi_F < \chi_I$ in order for the non-forecaster's IC constraint to bind. Thus when forecasting is expensive, the nature of the solution reverses: The supplier sacrifices efficiency with the forecaster to reduce the rents of the non-forecaster. Because a market with a forecaster is inherently more profitable, distorting the forecaster's actions may simply be too costly. The supplier may instead forego screening and offer just one contract.

4.3 Quantity flexibility analysis

We now turn to screening with QF contracts. We begin with an analog to Theorem 15.

Theorem 16 *Suppose retailers* 0 *and* 1 *are in independent markets and face stochastic demands* X_0 *and* X_1*, respectively. The supplier lets both select from a pair of QF contracts* $\{w_1, d_1\}$ *and* $\{w_2, d_2\}$*. Assume* $w_1 > w_2$*. If for all QF contracts* $\{w, d\}$,

$$\frac{M_1}{M_0} \geq \frac{\Phi_1(y_1^*(1-d))}{\Phi_0(y_0^*(1-d))}, \tag{10}$$

where y_j^* *is retailer* j*'s optimal order given* $\{w, d\}$ *and* $M_j = 1 - \frac{1}{y_j^*}\int_{y_j^*(1-d)}^{y_j^*} \Phi_j(\xi)\, d\xi$, *then*

1. If retailer 1 is indifferent between the two contracts, retailer 0 prefers contract $\{w_1, d_1\}$.
2. If retailer 0 is indifferent between the two contracts, retailer 1 prefers contract $\{w_2, d_2\}$.

Unlike the buy back case, we require a more stringent condition then the star order to assure that the IC constraints can cross only once. This is due in part to the difficultly in comparing decision across markets and contracts. Retailers 0 and 1, for example, will not necessarily choose the same service level for a given contract, and it is not the case that $y_0^1 > y_0^2$ implies that $y_1^1 > y_1^2$. However, one can argue that condition (10) is more likely to hold when $X_1 \leq_* X_0$.[5] By Theorem 9, if X_1 is smaller in the star order then the right hand side of (10) is less than one. The value M_j may be interpreted as the expected fraction of units in the supply chain for which the retailer ultimately pays. There is reason to believe

[4] This requires $\varepsilon = (r - c)(\tilde{\tau} + \tilde{\kappa})/\pi_1^I$.

[5] If X_0 and X_1 have power function distributions, $X_1 \leq_* X_0$ implies condition (10).

that this will be higher in a less variable market. For example, if X_1 is smaller in the star order, then retailer 1 is more likely than retailer 0 to sell more than his minimum purchase. The condition also appears to hold for the examples we consider below.

Theorem 17 *Let $\{w_{d_N}, d_N\}$ be a coordinating QF contract as defined in Theorem 6 for the non-forecaster's distribution such that $\Pi_N^R(w_{d_N}, d_N) = \tau$. Suppose condition (10) holds. If $\Pi_F^R(w_{d_N}, d_N) \leq \pi_1^I$, then there exists a QF contract $\{w_{d_F}, d_F\}$ that is a coordinating contract for the forecaster's distribution such that $w_{d_F} < w_{d_N}$ and $d_F < d_N$ and $\{w_{d_N}, d_N\}$ and $\{w_{d_F}, d_F\}$ together satisfy the constraints of $\tilde{\mathbf{P}}\mathbf{1}$.*

The theorem illustrates that there is something concrete that can be accomplished with coordinating QF contracts that cannot be accomplished with coordinating buy back contracts. The supplier can screen forecasting from non-forecasting retailers because the coordinating QF contracts depend on the demand distribution. Thus while dependence on the demand distribution may initially appear to be a drawback, here it has a benefit: The supplier can tell forecasters from non-forecasters without sacrificing any supply chain profits.

Although the supplier does not necessarily have to sacrifice supply chain profits, she may in fact do so. Consider the case of $\tilde{\kappa}_F = 0$. Because $\tilde{\Pi}_F^R(w_{d_N}, d_N) > \tilde{\tau}$, the supplier can earn more in the event that the retailer can forecast if she reduces the attractiveness of the non-forecaster's contract. The analysis is similar to the buy back case. By inducing the non-forecaster to serve a greater fraction of demand, the supplier reduces the rents the forecaster earns. Her ability to do this is again potentially limited by a positive cost of forecasting.

4.4 Relative performance buy back and QF contracts

We now examine the relative performance of QF and buy back contracts through a numerical example. We consider the case of exponential demand (i.e., $k = 1$ in (5)) because, given a gamma prior, it leads to relatively simple expressions for profit functions (Lariviere and Porteus, 1999). For example, the normalized return for the centralized system is:

$$\tilde{\pi}_a^I = r - c - ac\left(\left(\frac{r}{c}\right)^{1/a} - 1\right).$$

We set $r = 10$ and $c = 2.5$, yielding a critical fractile of 75%. The parameters of the gamma prior are $(S_0, a_0) = (100, 1.1)$, giving a mean demand of $1,000$. Forecasting is modeled as seeing five independent demand draws. The posterior shape parameter is then $a_1 = a_0 + n = 1.1 + 5 = 6.1$. The probability that the retailer can afford to forecast (i.e. ρ) is 50%. With this parameter set, forecasting has a significant impact on supply chain profitability. The normalized return for a non-forecasting supply chain $\tilde{\pi}^I_{a_0}$ is 0.552 while $\tilde{\pi}^I_{a_1}$ is 3.609. The normalized return for an exponential distribution would be $\lim_{a\to\infty} \tilde{\pi}^I_a = 4.034$. Below we vary the retailer opportunity cost τ and the forecaster's cost κ_F to determine their impact on which contractual form performs best.

We first consider the case of very cheap forecasting, i.e., $\kappa_F = 0$. We begin by restricting the supplier to offering coordinating contracts. In light of Theorems 14 and 17, one might anticipate that QF contracts will outperform buy backs in this setting. After all, the latter cannot distinguish between the high and low types while the former can. Consequently, with buy backs the supplier is only offering one contract while with QF contracts he is offering a menu of contracts. The results in Table 1 are therefore somewhat surprising. Here we report the expected profit of the supplier (averaged over realizations of a forecasting and a non-forecasting retailer) as well as each type of retailer for various values of τ. Profits are gross of opportunity and forecasting costs and are reported for a mean demand of $1,000$. The retailer's opportunity cost is varied from 20% to 80% of $\pi^I_{a_0}$. We see that the supplier is better off using buy backs in all cases. For a low retailer opportunity costs, the difference is relatively small but for higher costs moving from QF to buy backs would increase the supplier's profit by five to ten percent.

As seen in Table 2, this difference does not go away as one moves to optimal contracts. Moving from QF to buy back contracts again significantly increases supplier profit when participation costs are high. We do not report non-forecaster profit because the non-forecaster will be driven to indifference. Rather we focus on the steps the supplier takes to limit the forecaster's profit. We see that the supplier does indeed induce the non-forecasting retailer to serve a greater fraction of demand and thus sacrifices some channel efficiency in the event the retailer does not forecast. This is accomplished by offering the non-forecaster greater flexibility in returning the product while offering the forecaster a cheaper wholesale price. Efficiency losses can be considerable. Pushing the service level above 85% lowers the efficiency of the supply chain (i.e., the fraction of possible profit captured) to around 70%. Foregone profit from the non-forecaster are only partially offset by rents taken from the forecaster. For example, when τ equals 60% of $\pi^I_{a_0}$, moving from a coordinating QF contract to an optimal one reduces the forecasters profits by over \$240 but the supplier's expected profit increases by less than \$70. We infer that the supplier is taking a significant reduction in profit when the retailer cannot forecast. In fact, the supplier has an expected loss when the retailer cannot forecast and τ equals 80% of $\pi^I_{a_0}$. This is true under both buy back and QF contracts. In this case, she would prefer to offer a contract that only the forecaster would accept and forego trading with the non-forecater.

Table 3 offers some insight into why buy backs perform better than QF contracts. Here we report the retailer's expected share of system profit when the retailer forecasts. Recall that the general structure of the solution is to force the non-forecaster's individual rationality constraint to bind and then offer the forecaster a contract that leaves him indifferent between the two available contracts. Table 3 shows that the forecaster simply finds the non-forecaster's QF more attractive than the non-forecaster's buy back contract. Consider first the case of coordinating contracts. Coordinating buy backs cannot separate the types of retailers but they do cap what the forecaster earns. If the non-forecaster captures, say, 40% of system profit, the forecaster only captures 40% as well. Such is not the case with QF contracts. If the forecaster were to take the non-forecaster's contract, he would earn more than 40% of system profit. This asymmetry carries over to the optimal contracts as well.

Theorem 9 offers some intuition for this. Given the same QF contract, the retailer in the smaller market (in the sense of the star order) chooses a higher service level. Consequently, the service level the forecaster picks under the non-forecaster's QF contract must be greater than the service level the non-forecaster picks under the same contract. In contrast, the forecaster picks the same service level as the non-forecaster when buy backs are used. It must be the case that the forecaster faces a lower marginal cost of increasing the service level when QF contracts are used. That is, the non-forecaster's QF contract delivers "cheap" stock to the forecaster, and any contract that forecaster would prefer to the non-forecaster's must equal this windfall profit.

Table 3 does, however, suggest a setting in which QF contracts may perform better than buy backs: when forecasting is expensive. From Tables 1 and 2, we see that there is a large range of forecasting costs over which the supplier will be able to implement her preferred contract. For QF contracts to be preferable, forecasting must be extremely expensive. We present such an example in Table 4. Here we continue to give τ as a percent of $\pi^I_{a_0}$ but give $\tau + \kappa_F$ as a percent of $\pi^I_{a_1}$. We see that both coordinating and optimal buy backs perform worse than coordinating QF contracts. (Using optimal QF contracts would increase supplier profit by less than 0.25%.) Note that under buy backs it is optimal for the supplier to use only one contract and forego separating forecasters from non-forecasters. QF contracts work well in this setting since the structure of the solution is the reverse of the cheap forecasting case. Now the forecaster's IR constraint binds, and it is the non-forecaster who is indifferent between the contracts. If the non-forecaster takes the forecaster's contract, he must lower the service level he would provide despite the lower wholesale price. It is thus easier for supplier to induce him to take a different contract.

This example is representative of several that we have examined, but it is worth noting several features that impact the results. First, and perhaps most importantly, forecasting dramatically boosts returns because (a) the system starts with a very diffuse prior and (b) the forecasting technology provides for a relatively large number of draws. There are decreasing returns in the number of draws. Hence, point (a) is more important. For example, if the forecasting technology provided for only three draws (i.e., $n = 3$) and the initial shape parameter a_0 remained 1.1, the supplier would still offer contracts that distort the non-forecasting retailer's actions significantly (although less than in the current example). However, if $n = 3$ and $a_0 = 3.1$ (so a_1 still equals 6.1), we would see relatively little distortion in the non-forecaster's actions under the optimal contract since forecasting would deliver a much lower increase in profit.[6]

A second consideration is the critical fractile of the integrated system. Beginning from a service level of 75% leaves room for a large increase. Obviously, if we had started from a level of, say, 95%, it would be hard to increase the service by, say, 10%. Finally, the prior on the ability of the retailer to forecast (i.e., ρ) matters. The distortion in the non-forecaster's action is increasing in ρ. Indeed, if we pick ρ sufficiently low or high, the supplier will offer only one contract. If ρ is near zero and forecasting is cheap, she will offer a contract that is acceptable to both with little or no distortion in the non-forecaster's actions. If ρ is near one, she will offer a contract that only the forecaster will accept.

5 Discussion

We have presented a model in which a supplier must offer contracts to a retailer who may

[6] $\tilde{\pi}^I_{3.1} = 3.130$ and $\tilde{\pi}^I_{4.1} = 3.376$, so moving from $\tilde{\pi}^I_{1.1} = 0.552$ to $\tilde{\pi}^I_{4.1}$ remains a dramatic increase while moving from $\tilde{\pi}^I_{3.1}$ to $\tilde{\pi}^I_{6.1} = 3.609$ is a more modest gain.

or may not be capable of gathering additional information. We restrict the supplier to two contractual forms previously studied in the literature, buy backs and quantity flexibility contracts. It is well know that coordinating buy back contracts are independent of the demand distribution. We show that coordinating QF contracts depend on the variability of the demand distribution as measured by the star order. A market that is smaller in the star order requires a higher wholesale price for a given level of flexibility.

The relation between coordinating contracts and the demand distribution plays a role in the supplier's ability to screen retailers who are capable of forecasting from those who are not. Coordinating buy backs cannot be used to differentiate between forecasters and non-forecasters while coordinating QF contracts can. Designing terms that do induce forecasters to take distinct contracts from non-forecasters using either buy backs or QF contracts requires restricted returns. Forecasters get a cheaper wholesale price but less generous returns while non-forecasters pay a higher price but have greater flexibility in returning stock.

Which contract form produces a greater profit for the supplier depends on retailer costs. If forecasting is inexpensive, buy backs do better. Under QF contracts, a forecaster will choose a higher service level than a non-forecaster facing the same contract. If the forecaster were to take the contract meant for the non-forecaster, he is able to buy additional stock at what he perceives to be a low expected cost. This results in a windfall profit. Under a buy back contract, on the other hand, both types of retailers choose the same service level. The forecaster consequently does not receive a boon of cheap inventory by switching to the non-forecasters contract. When forecasting is very expensive, the nature of the solution changes, and QF contracts perform better. The non-forecaster places a sufficiently high value on flexibility that he does not gain much from switching to the forecaster's contract.

Like all models, the one presented here has limitations. The two most obvious are considering only two contract forms and the model of forecasting employed. To address the former, we chose to focus on buy backs and QF because they are simple contracts that have been studied in the literature but not directly compared. In addition, they provide a clean distinction between flexibility based on price and flexibility based on quantity. Clearly, the supplier could do no worse if she were free to use one of each (e.g., offering the non-forecaster a buy back contract but the forecaster a QF contract). Also, one could consider more complex contracts that offer return rates based on the fraction of the order returned. This alternative would likely better enable the supplier to tailor contracts to the demand distributions that the different types face. However, it would complicate the analysis significantly. For example, the retailer's ordering policy could potentially be quite complex.

Here we have used a very specific model of forecasting: a Bayesian inventory approach with Weibull demand and a gamma prior on the unknown rate. The primary advantage of this model is that it allows for a state space reduction so that the optimal stocking level after forecasting and expected profit from forecasting depend on the realized market signal in a very simple way. This allows us to ignore the exact outcome of forecasting in designing contracts. The Weibull-gamma pair is not the only combination of distributions for which this will hold. One could also consider gamma demand with a gamma prior or uniform demand with a Pareto prior (Azoury, 1985).

That said, many of our results do not rely on the specifics of the forecasting model. For example, that coordinating buy backs cannot achieve separation while QF contracts can will hold in general. Additionally, achieving separation with buy backs by offering restricted returns to the forecaster (Theorem 15), depends only on the star ordering. The Weibull-gamma assumption, however, is key to the numerical evaluation of contract performance. Using an alternatively model of information here (e.g., a bivariate normal) would be difficult since (for at least the QF case) the best contract (whether coordinating or otherwise) would depend on knowing the exact realization of the forecast.

There are a number of possible ways to generalize the model. One is to make the forecasting technology endogenous. Suppose all retailers have the same cost of forecasting which is increasing in the number of draws seen. One could examine which contracts induce the appropriate level of forecasting effort. Intuition suggests that a decentralized supply chain would collect less data than a centralized one even if the ex post stocking quantity is efficient. A second intriguing generalization is to consider multiple retailers. Suppose that each retailer is a local monopolist but that each faces the same demand distribution. Realized demands draws are independent across markets. If each forecasts and truthfully reveals the outcome of that forecast to the supplier, the supplier is much better informed than any one retailer. Issues to consider include how many retailers should forecast as well as retailer free riding.

Appendix A. Proofs

Proof of Lemma 1: Consider retailer profit given stocking level y and realized demand ξ :

$$g(y,\xi) = -wy^y + r\min\{y,\xi\} + b\max\{y-\xi,0\}.$$

$g(y^0,\xi)$ is piecewise linear and concave in ξ. Hence, $E[g(y^0,X_0)] \leq E[g(y^0,X_1)]$, but $E[g(y^0,X_0)] = \Pi_R^0$ and $E[g(y^0,X_1)] \leq E[g(y^1,X_1)] = \Pi_R^1$, by the optimality of y^1.

For the second part, let Φ_i denote the distribution of X_i. $\Phi_1(\xi) = \Phi_0(\xi/\theta)$ and $\phi_1(\xi) = \frac{1}{\theta}\phi_1(\xi/\theta)$. $y^1 = \theta y^0$ follows immediately. Next, using a change of variables, we have:

$$\int_A^B \xi\phi_1(\xi)\,d\xi = \int_A^B \frac{\xi}{\theta}\phi_1(\xi/\theta)\,d\xi = \theta\int_{A/\theta}^{B/\theta} z\phi_0(z)\,dz,$$

which together with $y^1 = \theta y^0$ implies that $\Pi_R^1 = \theta\Pi_R^0$ and $\Pi_S^1 = \theta\Pi_0^1$. ■

Proof of Lemma 2: Let y^i be the retailer's optimal order under contract i for $i = 1,2$. First, consider the case in which $b_1 \leq b_2$. Clearly the retailer prefers contract 2 since it offers a cheaper price and greater flexibility. Now suppose that $\frac{r-w_1}{r-b_1} = \frac{r-w_2}{r-b_2}$ so $y^1 = y^2$. Write the optimal expected profit under contract i as

$$\Pi_R^*(w_i,b_i) = (r-b_i)\int_0^{y^i} \xi\phi(\xi)\,d\xi.$$

Thus, $\Pi_R^*(w_2,b_2) - \Pi_R^*(w_1,b_1) > 0$ and the retailer prefers contract 2. Finally suppose that $b_1 > b_2$ but $\frac{r-w_1}{r-b_1} < \frac{r-w_2}{r-b_2}$. Then there exists a $\hat{b}$ such that $b_2 > \hat{b}$ and $\frac{r-w_1}{r-b_1} = \frac{r-w_2}{r-\hat{b}}$. The retailer prefers $\left\{w_2,\hat{b}\right\}$ to contract 1 and contract 2 to $\left\{w_2,\hat{b}\right\}$. ■

Proof of Lemma 4: The proof is similar to that of Lemma 1. ■

Proof of Lemma 5: Obviously, if $d_1 \leq d_2$, the retailer prefers contract 2. Now suppose $y^1 = y^2$, which implies $d_1 > d_2$. We have:

$$\begin{aligned}\Pi_R^*(w_i,d_i) &= y^i\left((r-w_i)\bar{\Phi}\left(y^i\right) - w_i(1-d_i)\Phi\left(y^i(1-d_i)\right)\right)\\ &\quad + r\int_0^{y^i}\xi\phi(\xi)\,d\xi - w_i\int_{y^i(1-d_i)}^{y^i}\xi\phi(\xi)\,d\xi = r\int_0^{y^i}\xi\phi(\xi)\,d\xi - w_i\int_{y^i(1-d_i)}^{y^i}\xi\phi(\xi)\,d\xi,\end{aligned}$$

where the second equality follows from (4). Consequently,

$$\Pi_R^*(w_2,d_2) - \Pi_R^*(w_1,d_1) = w_1\int_{y^1(1-d_1)}^{y^1}\xi\phi(\xi)\,d\xi - w_2\int_{y^2(1-d_2)}^{y^2}\xi\phi(\xi)\,d\xi > 0.$$

The remainder of the proof is similar to that of Lemma 2. ■

Proof of Theorem 7: The first part is Theorem 3.C.1 in Shaked and Shanthikumar (1994). It is equivalent to $Z_1 \leq_{disp} Z_0$, where the dispersive order is as defined in footnote 3 and

$Z_i = \log X_i$. Let $\hat{\Phi}_i$ be the distribution of Z_i. $\hat{\Phi}_i(\xi) = \Phi_i\left(e^{\xi}\right)$, $\hat{\phi}_i(\xi) = e^{\xi}\phi_i\left(e^{\xi}\right)$, and $\hat{\Phi}_i^{-1}(\xi) = \log \Phi_i^{-1}(\xi)$. From Shaked and Shanthikumar (1994), $Z_1 \leq_{disp} Z_0$ if and only if

1. $\hat{\Phi}_0\left(\hat{\Phi}_0^{-1}(\alpha) + \lambda\right) \geq [\leq] \hat{\Phi}_1\left(\hat{\Phi}_1^{-1}(\alpha) + \lambda\right)$ for $\alpha \in (0,1)$ and $\lambda < [>] 0$ or
2. $\hat{\phi}_0\left(\hat{\Phi}_0^{-1}(\alpha)\right) \leq \hat{\phi}_1\left(\hat{\Phi}_1^{-1}(\alpha)\right)$ for $\alpha \in (0,1)$.

The first [second] alternative is equivalent to part 2 [3]of the theorem. ■

Proof of Theorem 8: Part 1 is Theorem 3.C.4 in Shaked and Shanthikumar (1994). For part 2, the inverse of $\theta_i X_i$ is $\theta_i \Phi_i^{-1}(\alpha)$. The result follows from Theorem 7 part 1. Part 3 follows from result 5.5 (a) in Barlow and Proschan (1975) and Theorem 7 part 1. ■

Proof of Theorem 9: Suppose the retailer chose the same service level β^*, i.e., $y_j = \Phi_j^{-1}(\beta^*)$. Equation (4) would imply that $\Phi_1(y_1(1-d)) = \Phi_0(y_0(1-d))$, but that cannot hold by Theorem 7 part 2. For the second part, manipulate (4) to yield:

$$\Phi_j(y_j(1-d)) = \frac{(r-w)}{w(1-d)}\bar{\Phi}_j\left(y^j\right).$$

The result follows from the first part of the theorem. ■

Proof of Theorem 10: Let $\beta^* = \frac{r-c}{r}$ be the critical fractile of the integrated channel. $w_d^j = c/\left(\frac{c}{r} + (1-d)\Phi_i\left((1-d)\Phi_j^{-1}(\beta^*)\right)\right)$. $w_d^1 \geq w_d^0$ by Theorem 7 part 2. The second part follows from Lemma 4. ■

Proof of Lemma 11: In proving the first part of the lemma, we drop the subscript i. Using the change of variables $t = \xi/S^{1/k}$, the j^{th} moment may be written as follows:

$$\int_0^\infty \xi^j \phi(\xi|S,a)\,d\xi = akS^{j/k}\int_0^\infty t^{n+k-1}\left(1+t^k\right)^{-(a+1)}dt.$$

Employing another change of variables – $u = \left(1+t^k\right)^{-1}$ – then yields:

$$akS^{j/k}\int_0^1 (1-u)^{j/k}u^{a-j/k-1}du = aS^{j/k}\frac{\Gamma(a-j/k)\Gamma(1+j/k)}{\Gamma(a+1)},$$

where the final equality follows from the definition of the beta function (Mood, Graybill, and Boes, 1974).

For the second part, note that $\Phi_i(\xi|S_i,a_i) = 1 - S_i^{a_i}\left(S_i+\xi^k\right)^{-a_i}$. Let $a_i = a$. It is straightforward to verify that $\Phi_0\left(S_0^{1/k}\xi/S_1^{1/k}|S_0,a\right) = \Phi_1(\xi|S_1,a)$. Given that this distribution is scaled in S, Theorem 8 implies that to prove the third part we only need to consider the case of $S_0 = S_1 = 1$. For simplicity, let $a_0 = a$ and $a_1 = a_0 + \delta$ for $\delta \geq 0$. We then have:

$$G(\xi) = \Phi_0^{-1}(\Phi_1(\xi))/\xi = \frac{\left(\left(1+\xi^k\right)^{1+\delta/a} - 1\right)^{1/k}}{\xi},$$

and

$$G'(\xi) = \frac{\left(\left(1+\xi^k\right)^{1+\delta/a} - 1\right)^{1/k}}{\xi^2 a \left(\left(1+\xi^k\right)^{1+\delta/a} - 1\right)} A(\xi, \delta).$$

where $A(\xi,\delta) = a + \left(1+\xi^k\right)^{\delta/a}\left(\delta\xi^k - a\right)$. Clearly, the sign of G' is the same as that of $A(\xi,\delta)$. Note that $A(\xi,0) = 0$. Further,

$$\frac{\partial A}{\partial \delta} = \left(1+\xi^k\right)^{\delta/a}\left(\xi^k - \ln\left[1+\xi^k\right]\left(1 - \delta/a\xi^k\right)\right) \geq 0,$$

since $\xi^k \geq \ln\left[1+\xi^k\right]$. $G(\xi)$ is consequently increasing and $X_1 \leq_* X_0$. ■

Proof of Theorem 12: First, note that $\tilde{X}_a \leq_{cx} \tilde{X}_{a'}$ for $a \geq a'$ by Theorem 8 part 1. The first two parts then follow from Lemma 1. The third part follows from Lemma 4. ■

Proof of Lemma 13: We begin with the case of $n = 1$ so forecasting yields only one observation. Consider the expectation of $S_1^{1/k}$:

$$\begin{aligned}
\int_0^\infty S_1^{1/k}\phi(\xi|S_0,a_0)\,d\xi &= \int_0^\infty \left(S_0+\xi^k\right)^{1/k}\phi(\xi|S_0,a_0)\,d\xi \\
&= \frac{a_0 S_0^{1/k}}{a_0 - 1/k}\int_0^\infty \phi(\xi|S_0,a_0-1/k)\,d\xi \\
&= \frac{a_0 S_0^{1/k}}{a_0 - 1/k}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
E[\mu(S_1,a_1)] &= E\left[S_1^{1/k}\right]\frac{a_1\Gamma(a_1-1/k)\Gamma(1+1/k)}{\Gamma(a_1+1)} \\
&= \frac{a_0 S_0^{1/k}}{a_0-1/k}\frac{(a_0+1)\Gamma(a_0+1-1/k)\Gamma(1+1/k)}{\Gamma(a_0+2)} \\
&= \frac{a_0 S_0^{1/k}\Gamma(a_0-1/k)\Gamma(1+1/k)}{\Gamma(a_0+1)} \\
&= \mu(S_0,a_0).
\end{aligned}$$

The penultimate equality follows from the fact that $\Gamma(a+1) = a\Gamma(a)$. For the case of $n > 1$, an induction yields

$$E\left[S_1^{1/k}\right] = \frac{\prod_{j=0}^{n-1}(a_0+j)}{\prod_{j=0}^{n-1}(a_0+j-1/k)}S^{1/k},$$

which leads to the desired result. ■

Proof of Theorem 14: Suppose a pair of contracts existed such that $\varepsilon_F > \varepsilon_N$ and all constraints are satisfied. (The case for $\varepsilon_F < \varepsilon_N$ is similar.) By Theorem 3, $\tilde{\Pi}_F^R(w_{\varepsilon_N}, b_{\varepsilon_N}) = \left(1 - \frac{\varepsilon_N}{r-c}\right)\tilde{\pi}_{a_1}^I > \left(1 - \frac{\varepsilon_F}{r-c}\right)\tilde{\pi}_{a_1}^I = \tilde{\Pi}_F^R(w_{\varepsilon_F}, b_{\varepsilon_F})$. Hence the forecaster's incentive compatibility cannot be satisfied unless $\varepsilon_F = \varepsilon_N$. ■

Proof of Theorem 15: Let Φ_j denote the distribution for random variable X_j. Given contract i, retailer j's optimal expected profit can be written as follows:

$$(r-b_i)\int_0^{\Phi_j^{-1}(\chi_i)} \xi\phi_j(\xi)\,d\xi = (r-b_i)\int_0^{\chi_i} \Phi_j^{-1}(\alpha)\,d\alpha,$$

where we have used the change of variables $\alpha = F(\xi)$. For retailer j to prefer $\{w_1, b_1\}$, it must be case that

$$(r-b_1)\int_{\chi_2}^{\chi_1} \Phi_j^{-1}(\alpha)\,d\alpha \geq (b_1-b_2)\int_0^{\chi_2} \Phi_j^{-1}(\alpha)\,d\alpha. \tag{A-1}$$

If retailer j is indifferent, (A-1) holds as an equality. Suppose X_0 and X_1 are scaled so that $\Phi_0^{-1}(\chi_2) = \Phi_1^{-1}(\chi_2)$. Then by Theorem 8 part 3,

$$\int_{\chi_2}^{\chi_1} \Phi_0^{-1}(\alpha)\,d\alpha \geq \int_{\chi_2}^{\chi_1} \Phi_1^{-1}(\alpha)\,d\alpha$$

and

$$\int_0^{\chi_2} \Phi_1^{-1}(\alpha)\,d\alpha \geq \int_0^{\chi_2} \Phi_0^{-1}(\alpha)\,d\alpha.$$

Thus for this scaling, if retailer 1 is indifferent, retailer 0 prefers contract 1. If retailer 0 is indifferent, retailer 1 prefers contract 2. By Lemma 1 and Theorem 8, if the results hold for this scaling, they hold for all scalings. ■

Proof of Theorem 16: Let $\{w, d_j(w)\}$ be the set of contracts that holds retailer j's profit constant at, say, $\bar{\Pi}$. Implicit differentiation and the envelope theorem (de la Fuente, 2000) leads to:

$$\begin{aligned} d_j'(w) &= -\frac{\partial \Pi_j^R/\partial w}{\partial \Pi_j^R/\partial d} = \frac{\int_{y_j^*(1-d)}^{y_j^*} \xi\phi_j(\xi)\,d\xi + y_j^*\bar{\Phi}_j\left(y_j^*\right) + y_j^*(1-d)\,\Phi_j(\xi)\,d\xi}{w y_j^*\Phi_j(\xi)} \\ &= \frac{M_j}{w\Phi_j(\xi)}, \end{aligned}$$

where the last equality follows from an integration by parts. Condition (10) then implies that $d_1'(w) \geq d_0'(w)$ for all w. Consequently, the iso-profit curves of retailer 1 in (w,d)-space are always steeper than those of retailer 0, which leads to the stated conclusions. ■

Proof of Theorem 17: Since $\Pi_F^R(w_{d_N}, d_N) \leq \pi_1^I$, there exists a coordinating QF contract $\{w_{d_F}, d_F\}$ such that $\Pi_F^R(w_{d_F}, d_F) = \Pi_F^R(w_{d_N}, d_N)$. The forecaster chooses a lower service level under $\{w_{d_F}, d_F\}$. Hence by Lemma 5, $w_{d_F} < w_{d_N}$ and thus $d_F < d_N$. Satisfying the constraints of of $\tilde{\mathbf{P}}\mathbf{1}$ then follows from Theorem 16. ■

References

Azoury, K. S., "Bayes Solution to Dynamic Inventory Models under Unknown Demand Distribution," *Management Science*, 31, 9 1985, 1150-1160.

Barlow, R. E., and F. Proschan, *Statistical Theory of Reliability and Life Testing Probability Models*, New York: Holt, Rinehart, and Winston, 1975.

Burr, I. W., "Cumulative Frequency Functions," *Annals of Mathematical Statistics*, 13, 2, 1942, 215-232.

Cachon, G. P., "Supply Chain Coordination with Contracts," working paper, University of Pennsylvania, 2001.

Cachon, G. P. and M. A. Lariviere, "Contracting to Assure Supply: How to Share Demand Forecasts in a Supply Chain," *Management Science*, 47, 5, 2001a, 629-646.

Cachon, G. P. and M. A. Lariviere, "Supply Chain Coordination with Revenue-Sharing Contracts: Strengths and Limitations," working paper, Northwestern University, 2001b.

Chen, F., "Information Sharing and Supply Chain Coordination," working paper, Columbia University, 2001a.

Chen, F., "Salesforce Incentive, Market Information, and Production/Inventory Planning," working paper, Columbia University, 2001b.

Corbett, C. J., "Stochastic Inventory Systems In A Supply Chain With Asymmetric Information: Cycle Stocks, Safety Stocks, and Consignment Stock," *Operations Research*, 49, 4, 2001, 487-500.

Corbett, C. J., and X. de Groote, "A Supplier's Optimal Quantity Discount Policy Under Asymmetric Information," *Management Science*, 46, 3, 2000, 444-450.

de la Fuente, A., *Mathematical Method and Models for Economists*, Cambridge University Press, Cambridge, UK, 2000.

DeGroot, M. H., *Optimal Statistical Decisions*, McGraw-Hill, New York, 1970.

Donohue, K. L., "Efficient Supply Contracts for Fashion Goods with Forecast Updating and Two Production Modes," *Management Science*, 46, 11, 2000, 1397-1411.

Gerchak, Y. and D. Mossman, "On the Effect of Demand Randomness on Inventories and Costs," *Operations Research*,42, 4, 1992, 804-807.

Ha, A., "Supplier-Buyer Contracting: Asymmetric Cost Information and Cut-off Level Policy for Buyer Participation," *Naval Research Logistics*, 48, 1, 2001, 41-64.

Lariviere, M. A., "Supply Chain Contracting and Coordination with Stochastic Demand," in S. Tayur, M. Magazine, and R. Ganeshan (Eds.), *Quantitative Models of Supply Chain Management*, Kluwer Academic Publishers, Boston, MA, 1999.

Lariviere, M. A., and E. L. Porteus, "Stocking Information: Bayesian Inventory Management with Unobserved Lost Sales," *Management Science*, 45, 3, 1999, 346-363.

Lariviere, M. A., and E. L. Porteus, "Selling to a Newsvendor: An Analysis of Price-Only Contracts," *Manufacturing & Service Operations Management*, 3, 4, 2001, 293-305.

Mood, A. M., F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, McGraw-Hill, New York, 1974.

Pasternack, B. A., "Optimal Pricing and Return Policies for Perishable Commodities," *Marketing Science*, 4, 2, 1985, 166-176.

Porteus, E. L., and S. Whang, "Supply Chain Contracting: Non-Recurring Engineering Charges, Minimum Order Quantity, and Boilerplate Contracts," working paper, Stanford University, 1999.

Scarf, H. E., "Bayes Solution to the Statistical Inventory Problem," *Annals of Mathematical Statistics*, 30, 2, 1959, 490-508.

Scarf, H. E., "Some Remarks on Bayes Solution to the Inventory Problem," *Naval Research Logistics Quarterly*, 7, 591-596.

Salanié, Bernard, *The Economics of Contracts*, The MIT Press, Cambridge, MA, 1997.

Shaked, M. and J. G. Shanthikumar, *Stochastic Orders and Their Applications*, Academic Press, San Diego, CA, 1994.

Tsay, A. A., "The Quantity Flexibility Contract and Supplier-Customer Incentives," *Management Science*, 45, 10, 1999, 1339-1358.

Tsay, A. A. and W. S. Lovejoy, "Quantity Flexibility Contracts and Supply Chain Performance," *Manufacturing and Service Operations Management* 1, 2, 1999, 89-111.

Zarley, C., "IBM PC Co. Tries to Determine Real Demand: Want ThinkPads? Assure no returns," *Computer Reseller News*, October 31, 1994, 1.

Table 1
Supply Chain Performance under Coordinating Contracts

τ (% of π^I_{a0})	*Buy Backs* Expected Supplier Profit	*Buy Backs* Expected Forecaster Profit	*Buy Backs* Expected Non-Forecaster Profit	*Quantity Flexibility* Expected Supplier Profit	*Quantity Flexibility* Expected Forecaster Profit	*Quantity Flexibility* Expected Non-Forecaster Profit
20%	1,664.54	721.77	110.50	1,641.09	768.66	110.50
40%	1,248.40	1,443.54	221.00	1,209.16	1,522.01	221.00
60%	832.27	2,165.31	331.50	785.99	2,257.86	331.50
80%	416.13	2,887.07	442.00	376.71	2,965.93	442.00

Table 2
Supply Chain Performance under Optimal Contracts

Buy Backs

τ (% of π^I_{a0})	Expected Supplier Profit	Expected Forecaster Profit	Non-Forecaster Efficiency	Non-Forecaster Service Level	$\{w_F,b_F\}$	$\{w_N,b_N\}$
20%	1,673.03	689.55	97.2%	79.3%	{8.57,8.09}	{8.70,8.36}
40%	1,279.34	1,330.50	90.7%	82.2%	{7.23,6.31}	{7.66,7.15}
60%	896.47	1,937.78	82.1%	84.3%	{5.97,4.63}	{6.76,6.15}
80%	522.36	2,519.87	72.0%	85.9%	{4.76,3.02}	{5.94,5.27}

Quantity Flexibility

	Expected Supplier Profit	Expected Forecaster Profit	Non-Forecaster Efficiency	Non-Forecaster Service Level	$\{w_F,d_F\}$	$\{w_N,d_N\}$
20%	1,647.56	704.67	90.8%	82.2%	{8.63,0.828}	{8.98,0.920}
40%	1,243.70	1,390.24	88.7%	82.8%	{7.28,0.724}	{7.98,0.877}
60%	850.61	2,010.86	78.7%	84.9%	{6.10,0.626}	{7.31,0.875}
80%	468.96	2,602.83	67.7%	86.4%	{4.80,0.486}	{6.44,0.843}

Table 3
Fraction of Supply Chain Profit Captured by a Forecasting Retailer

τ (% of π^J_{a0})	Coordinating Buy Backs	Coordinating QF	Optimal Buy Backs	Optimal QF
20%	20.0%	21.3%	19.1%	19.5%
40%	40.0%	42.2%	36.9%	38.5%
60%	60.0%	62.6%	53.7%	55.7%
80%	80.0%	82.2%	69.8%	72.1%

Table 4
Supply Chain Performance with Expensive Forecasting

		Coordinating Buy Backs			*Optimal Buy Backs*				*Coordinating QF*		
τ (% of π^J_{a0})	$\tau + \kappa_F$ (% of π^J_{a1})	Expected Supplier Profit	Expected Forecaster Profit	Expected Non-Forecaster Profit	Expected Supplier Profit	Efficiency Given Forecaster	Efficiency Given Non-Forecaster	Service Level	Expected Supplier Profit	Expected Forecaster Profit	Expected Non-Forecaster Profit
20%	25%	1,560.50	902.21	138.12	1,560.58	99.997%	99.993%	74.8%	1,567.19	902.21	124.75
40%	45%	1,144.37	1,623.98	248.62	1,144.63	99.989%	99.978%	74.6%	1,154.05	1,623.98	229.26
60%	65%	728.23	2,345.75	359.12	728.78	99.977%	99.955%	74.4%	738.51	2,345.75	338.57
80%	85%	312.10	3,067.52	469.62	313.03	99.961%	99.923%	74.2%	319.27	3,067.52	455.28