Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Coordinating Replenishment and Pricing in a Firm

Qing Li, Derek Atkins,

Coordinating Replenishment and Pricing in a Firm

Qing Li • Derek Atkins
Department of Information and Systems Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
Faculty of Commerce and Business Administration, The University of British Columbia, 2053 Main Mall, Vancouver, British Columbia, Canada V6T 1Z2
imqli@ust.hk • derek.atkins@commerce.ubc.ca

Replenishment and pricing strategies are traditionally determined by entirely separate units of a firm, the former by production and the latter by marketing. In a large organization production and marketing are traditionally measured in terms of performance criteria appropriate and relevant within the world in which they operate, rather than in terms of overall company performance. We consider a situation where the headquarters uses a simple linear transfer price between the two functions to govern the transactions. The misaligned incentives among these functional managers are caused by a transfer price between them that distorts the marginal production cost and revenue, as well as by a misallocation of cost: Marketing's pricing strategy influences expected leftover inventory, but only production incurs the cost. The misalignment can be mitigated through the following two ways. First, if production commits to a service level instead of an inventory level, both production and marketing, and hence the firm as a whole, are better off. Second, the same improvement can be achieved through organizational changes so that marketing becomes the dominant function. We also propose a mechanism that aligns the functional managers' incentives to be compatible to the firm.

(*Inventory; Pricing; Coordination; Manufacturing/Marketing Interface*)

1. Introduction

For the majority of companies involved in some form of manufacturing or service delivery, two types of operating decisions are of great importance. One concerns the pricing of finished products and the other the replenishment of their inputs. For a company of at least a reasonable size it is common that these two clusters of decisions are delegated to disparate organizational units, often with few organizational tools to assist with their coordination. To take a typical situation: A production manager might be involved in the purchasing of raw materials, components, or even the setting up or purchasing of production capacity in advance of a selling season and would be making such decisions on the basis of demand forecasts and experience. Simultaneously across the organization, maybe geographically as well as organizationally remote, a marketing manager could be making decisions on prices for finished goods or services. Both will be working within a framework that evaluates their respective performance, and it is not unusual that such performance evaluations relate most directly to criteria easily comprehended and measured within their own units. Unless a well worked out scheme is in place, such local measures of departmental performance can lead to these two critical operating decisions being made in a way that is suboptimal to the firm as a whole.

To study this issue, we consider a highly simplified business scenario. Production chooses the amount of a single-inventory replenishment of a single-input material. For marketing the task at hand is to set the

1523-4614/02/0404/0241$05.00
1526-5498 electronic ISSN

retail price of a single finished product. Both choices must be made in the face of uncertainty about the upcoming season, with only a forecast of demand in the form of a final product demand distribution shared with the company as a whole. We simplify this world to have a single season, an approximation to either a seasonal product or any other situation in which the carry forward of either raw material inputs or finished goods is not the prime issue. This single-period assumption is primarily made to focus attention on the key aspects of the relationships between these two operating decisions and not to confound the issue with multiyear effects. The demand is price sensitive and random. This simple scenario captures two of the most generic functions of supply chains: price setting and inventory planning (Stern and El-Ansary 1992). The headquarters (HQ) of the firm has the objective of maximizing total expected profit and is thus motivated to find proper schemes to align the department managers' incentives. The purpose of this paper is to address coordination issues in such an environment. A simple procedure is for the HQ to make both production and marketing into profit centers by establishing a linear transfer price at which marketing must ''purchase'' product from production. As far as the interaction between production and marketing is concerned, the transfer price is exogenous. A question HQ might have is whether the linear transfer price creates misaligned incentives, and if so, what the consequences of this might be.

If we were to ignore the inventory decision and replace the unknown demand distribution with a deterministic demand function, then we would face a scenario within the well-studied realm of double marginalization; see, for example, Spengler (1950). In this scenario a transfer price that is higher than the marginal cost of production will lead to a higher retail price and a lower total profit than would be chosen in an integrated system. A much remarked upon ''solution'' to the problem of double marginalization, the *two-part tariff,* replaces the single proportional transfer price by a (affine) sharing rule: a transfer price and a fixed ''side-component.'' To make this work and coordinate the transaction so that the outcome is indistinguishable from an integrated system, the transfer price has to be set equal to the marginal cost of production, and production then simply makes a profit from the fixed portion of the fee. Alternatively, Jeuland and Shugan (1983) suggest a *profit sharing contract or quantity discount.* Cachon (1998) reviews a variety of models. One of the models raises the question of how many units of product to deliver to the market rather than what price should be charged with stochastic demand and marketing as a price taker. Although the decision variable is different, the result of this model has essentially the same interpretation as that of double marginalization: Production makes a profit by setting a wholesale price higher than the marginal cost of production, and marketing chooses a lower quantity than would be optimal for the system. Note that in all the work cited so far production has no manufacturing operations; in a sense production ''adds no value,'' and there is no incentive that is incompatible with the firm.

Although offering certain insights, these models based on double marginalization fail to capture key coordination issues. The principal reason is that production in these models adds nothing of value other than adding to the cost. We might contrast this to the classic integrated multiechelon model of Clark and Scarf (1960). Here each stage has to hedge against uncertainty during the leadtime and, in the case of fixed set-up costs, to carry inventory because of scale economies. When studying the decentralized counterpart to this setting, Lee and Whang (1999) need coordinating rules far more complex than a two-part tariff. In two extensions to Cachon (1998) by Emmons and Gilbert (1998) and Cachon and Lariviere (2000), both pricing and stocking are considered in a strategic setting. However, the upper echelon in their models still does not perform an operations function, so their results remain a variation of double marginalization. There is a considerable literature on the coordination of supply chains in which production does perform the function of production/inventory planning, but with the stochastic demand typically exogenously given. In this case there is no pricing role for marketing (e.g., Cachon and Zipkin 1999, Caldentey and Wein 2000). It is not our intention to give a comprehensive review here. Those interested should refer to Cachon (1998) and Anupindi and Bassok (1998).

Another stream of research focuses on the manufacturing/marketing interface. Different conflicts inherent in these two functions are addressed. For example, Porteus and Whang (1991) examine the effort levels in building capacity by the manufacturing manager and the sales effort by the product managers. De Groote (1994) studies the conflicts with product diversity. There is probably much more nonanalytic work on this subject. See Eliashberg and Steinberg (1993) for an early comprehensive review on both analytic and nonanalytic work. Recently, Chen (2000a) focused on the market information possessed by sales people. The firm must design an incentive scheme to motivate sales people not only to work hard, but also to truthfully disclose their information about the market because the information is important for the firm's production and inventory planning decisions. The paper studies a scheme proposed by previous research and compares it with menus of linear contracts. Proper pricing used by marketing can also help supply chain management. Chen (2000b) considers a monopolistic firm selling a single product to a market where the customers may be enticed to accept a shipping delay by being offered a discounted price. Customers are heterogeneous in terms of aversion to delays. When a customer agrees to wait, the firm gains advanced demand information that can be used for planning inventory. The paper shows how the pricing and replenishment decisions should be jointly made so that the costs, due to discounted prices, and the benefits, due to advanced demand information, are balanced. The papers cited above do not consider either pricing or coordination in a strategic setting.

Our research differs in that we focus on the incentive and coordination issues in a firm where replenishment and pricing decisions are made by production and marketing, respectively, in a decentralized fashion. We first examine the incentive misalignment between marketing and production when the HQ uses a linear transfer price between them. The misaligned incentives are caused by the transfer price between them distorting the marginal revenue and marginal cost facing production and marketing, respectively. In addition there is a misallocation of costs, because marketing's pricing strategy influences expected leftover inventory but only production incurs the cost. We consider two alternatives. The first is to develop a coordinating contract, but it is complex. The other alternative does not achieve coordination, but it allows HQ to maintain the simple linear transfer price and simply requires that production commit to a service level instead of an inventory level. We show that the latter can lead to improved performance for both production and marketing, and thus the firm as a whole. The two forms of commitment from production are equivalent operationally in an integrated system in which HQ jointly makes the replenishment and pricing decisions. We also study the issue of functional dominance and find that having marketing as the leader in the supply chain achieves the same benefit as having production commit to a service level.

The remainder of the paper is organized as follows: The next section describes the model in more detail. Section 3 investigates the integrated system. Sections 4 and 5 investigate the decentralized systems. We propose a scheme that can achieve coordination in §6 and give some numerical studies to compare the integrated systems and decentralized systems in §7. There are discussions on Stackelberg games in §8. Finally, conclusions and future research in §9 close the paper. Most of the proofs are delegated to the Appendix for convenience.

2. Model Description

The chronology of events, notation, and information structure are as follows:

(1) HQ sets a scheme to determine the performance measures for production and marketing. This includes choosing between a simple linear transfer price or a more sophisticated contract, and setting the rule to require production to commit to either a service level or an inventory level.

(2) Production and marketing, with a shared common forecast of demand represented by a random variable with a known distribution, make their decisions to maximize their expected profits according to

the scheme set by HQ. The marginal cost of production or ordering is c_k. The inventory here can be thought of quite generally. It could be the actual finished goods inventory level but would more likely be purchase orders of raw material, the reservation of production capacity in a third-party facility, the hiring of personnel, or other decisions that have to be made before the demand is known.[1]

(3) The random demand is observed.

(4) Marketing submits an order of the total realized demand to production, but production can only fulfill the smaller of the order and the preset inventory. There is a variable fulfillment cost c_p per unit charged to production, which includes all the variable costs borne in the logistic process, such as shipping, handling, and labeling.

Modeling the price-sensitive random demand is crucial in this research. We consider an additive demand of the following form: $D(p, \tilde{\epsilon}) = y(p) + \tilde{\epsilon}$, where $y(p)$ is a downward sloping demand curve that captures the dependency between demand and price and $\tilde{\epsilon}$ a random variable. We further define $y(p) = a - bp(a > 0, b > 0)$. The linear demand curve is common in the economics literature. Let $\Phi(\cdot)$, $\phi(\cdot)$, and $h(\cdot)$ be the cumulative distribution function, density function, and failure rate function of $\tilde{\epsilon}$, respectively. To make some of the analysis easier, we restrict ourselves to random variables with an increasing failure rate (IFR). This assumption is not too restrictive because it includes many common distributions. To avoid negative demand, we assume $\tilde{\epsilon}$ is defined in a certain range $[\alpha, \beta]$ and $y(c_k + c_p) + \alpha \geq 0$. This can easily be satisfied; for example, when both $\tilde{\epsilon}$ and the deterministic demand are nonnegative. Let μ be the mean of $\tilde{\epsilon}$.

There is considerable research that deals with simultaneously choosing an optimal stocking quantity and a price in the newsvendor setting with price-sensitive random demand. The integrated system in our model is similar to this problem. To our knowledge, Whitin (1955) was the first to formulate a newsvendor model with price effects. Porteus (1990) and Petruzzi and Dada (1999) are good sources of reviews for this problem. The latter also provides valuable new interpretations. Refer also to Federgruen and Heching (1999) for a review and recent results on the dynamic version of this problem. Starting from Ernst (1970), the joint inventory and pricing problem is often analyzed with a change of decision variable for convenience. Let K be the inventory the newsvendor wants to set. Define $z = K - y(p)$. The problem is transformed to maximizing the total expected profit over z and p. This transformation has been used as a mere mathematical convenience to analyze the price-sensitive newsvendor problem. Petruzzi and Dada (1999) recently provide its managerial interpretation: z can be interpreted as a safety stock factor. Note that there is a direct mapping from any given quantile of the demand distribution to a safety stock factor, but for a given inventory level K, the service level is dependent on the price.[2] To emphasize this distinction, we will call z a *service level*. This distinction will become important in the decentralized systems. We call a decentralized system in which production commits to a service level z a *service level game* and an *inventory game* if production commits to an inventory level K. We compare the outcomes of these two games.

Another dimension we consider is the issue of functional dominance in the firm. We study the Nash game when production and marketing have equal power, and they make the replenishment and pricing decision simultaneously. When there is one dominant function, the decisions will be made sequentially, and we will model this situation by a Stackelberg game.

Superscripting identifies the control scheme being considered. The superscript takes values of I to denote control by a single integrated decision maker (HQ) (§3), K the inventory game (§4), and Z the service level game (§5), respectively. There are two sets of subscripts. The subscript N refers to a solution

[1]The timing of the retail price determination is important in the sequel, and here we are most interested in when the price also must be set well in advance of knowledge of the market demand. There are situations where inventory (price) is more flexible than price (inventory); that is, inventory (price) is postponable. Van Mieghem and Dada (1999) give practical examples for all these cases.

[2]Service level here is defined as the probability of not stocking out (that is, the type 1 service level in Nahmias 1989). When the safety stock factor is given as z, the service level is $\Phi(z)$. When the inventory level K is given, the service level is $\Phi(K - y(p))$, a function of p.

from the Nash game, *PS* to the production Stackelberg game, and *MS* to the marketing Stackelberg game, respectively. The subscripts p and m are used to differentiate expected profit functions and reaction functions: p represents production and m represents marketing.

3. Integrated Systems

The integrated systems will be used as a benchmark against which to measure the performance of the decentralized systems. The integrated system is similar to the newsvendor problem with a price-sensitive demand. Let $G^I(p, K)$ be the total expected profit, then:

$$G^I(p, K) = -c_k K + (p - c_p) E \min(K, \tilde{D}) = I(p) - L(K, p) \tag{1}$$

where

$$I(p) = (y(p) + \mu)(p - c_k - c_p)$$

$$L(K, p) = c_k \Lambda(K - y) + (p - c_k - c_p)\Theta(K - y)$$

and where $\Lambda(x) = \int_\alpha^x (x - t)\phi(t)\, dt$ and $\Theta(x) = \int_x^\beta (t - x)\phi(t)\, dt$.

$I(p)$ represents the riskless profit function, the profit for a given price when the demand shock is replaced by its constant mean μ. Notice that without uncertainty in the demand side, production can precisely build the amount of inventory demanded. $L(K, p)$ is the loss function, which assesses an overage cost c_k for each unit of the expected unused inventory $\Lambda(K - y)$ if K was chosen too high and an underage cost $(p - c_k - c_p)$ for each unit of the $\Theta(K - y)$ expected shortages if K was chosen too low. Thus we solve:

$$\max_{p,K} G^I(p, K). \tag{2}$$

Define p^0 as the riskless price, the price that maximizes $I(p)$; then $p^0 = [a + \mu + b(c_k + c_p)]/(2b)$. The following lemma gives the optimal solution for the integrated system.

Lemma 1. *The optimal inventory and pricing policy is to build an inventory of $K^I = z^I + y(p^I)$ and to sell at the unit price p^I, where $p^I = p^0 - \Theta(z^I)/(2b)$ and z^I is uniquely given by $\Phi(z^I) = (p^I - c_k - c_p)/(p^I - c_p)$.*

Starting from Ernst (1970), this joint inventory and pricing problem is often analyzed with a change of decision variable for convenience. That is, the objective is set to maximize the expected profit over service level and price. The expected profit, expressed in terms of z and p and denoted as $\Pi^I(p, z)$, is

$$\Pi^I(p, z) = (y + \mu)(p - c_k - c_p) - (p - c_k - c_p)\Theta(z) - c_k\Lambda(z), \tag{3}$$

and the optimization problem is

$$\max_{p,z} \Pi^I(p, z). \tag{4}$$

It can be shown that this transformation does not sacrifice optimality. Formally,

Lemma 2. *(p^I, z^I) is the unique optimal solution of (4).*

Lemma 2 states that in an integrated system, jointly setting a service level and price is merely equivalent to jointly setting an inventory level and price. As we will see later, however, this distinction is crucial in decentralized systems: Whether production commits to an inventory level or a service level can lead to different performances for every party involved.

4. Decentralized Systems: Inventory Games

We investigate the consequences of HQ using a fixed transfer price w to make production and marketing profit centers. To guarantee production's participation, HQ will always set $w \geq c_p + c_k$. In this section, we assume that production's strategic variable is the inventory K, so the problem facing production is to find an inventory level that maximizes the expected profit for a given retail price while marketing chooses a price to maximize the expected profit, assuming that K is fixed. The transfer price w is exogenously given in the strategic interactions between production and marketing, reflecting the fact that functional managers in a firm are typically not involved in designing performance measures. Instead, they will try to maximize whatever performance measure that has been set by HQ. We refer to this decentralized system as an inventory game. That is, for a given p, production solves

$$\max_K \Pi_p^K(K) \tag{5}$$

where

$$\begin{aligned}\Pi_p^K(K) &= -c_k K + (w - c_p)E \min(K, \tilde{D}) \\ &= (w - c_k - c_p) \\ &\quad \times \left[(y + \mu) - \int_{K-y}^{\beta} (t - K + y)\phi(t)\, dt\right] \\ &\quad - c_k \int_{\alpha}^{K-y} (K - y - t)\phi(t)\, dt. \end{aligned} \tag{6}$$

Solving (5) yields the unique reaction function for a given price as

$$K_p(p) = y(p) + \Phi^{-1}[(w - c_k - c_p)/(w - c_p)]. \tag{7}$$

Marketing's profit function is

$$\Pi_m^K(p) = (p - w)\left[(y + \mu) - \int_{K-y}^{\beta} (t - K + y)\phi(t)\, dt\right],$$

and its unique reaction function $p_m^K(K)$ is given implicitly by

$$\begin{aligned}&-2b(p_m^K - p_m^0) - \Phi(K - y) \\ &\quad + b(p_m^K - w)\bar{\Phi}(K - y) = 0\end{aligned} \tag{8}$$

where $p_m^0 = (a + \mu + bw)/(2b)$, the price that maximizes $(p - w)(y + \mu)$.

THEOREM 1. *The inventory game has a unique Nash solution (K_N^K, p_N^K) given by (7) and (8).*

PROOF. The existence of a Nash solution is confirmed by the concavity of both expected payoff functions. To show uniqueness, it is sufficient to show that the derivative of one reaction curve is always strictly smaller than that of the other. Obviously, if we draw the curve with p as the vertical axis, then the reaction curve given by (7) has a slope of $-1/b$. The derivative of the reaction curve given by (8) is

$$\frac{dp_m^K}{dK} = \frac{\bar{\Phi}(K - y) - b(p_m^K - w)\phi(K - y)}{b^2(p_m^K - w)\phi(K - y) + 2b\Phi(K - y)} > -\frac{1}{b},$$

thereby proving the uniqueness of the Nash solution. □

The uniqueness relies on our assumption that the random shock $\tilde{\epsilon}$ has a probability density function. When the random shock is replaced by its mean value μ (that is, in the case of perfect information) the game may have multiple equilibria. See Li and Atkins (2001) for a detailed discussion on this subtlety and one possible way to overcome this frustration. Notice that $-1/b$ is the slope of the expected demand function. The inequality $dp_m^K/dK > -1/b$ has been termed a *price-smoothing effect* in the economics literature. See Amihud and Mendelson (1983) for a further discussion on this topic. In this context, the price response to the available inventory is milder than implied by the demand function.

Equation (6) can be rewritten as

$$\begin{aligned}\Pi_p^K(z) &= (w - c_k - c_p)E[\text{Sales}(K, p)] \\ &\quad - c_k E[\text{Unused inventory}(K, p)]\end{aligned} \tag{9}$$

and marketing's profit as

$$\Pi_m^K(p) = (p - w)E[\text{Sales}(K, p)] \tag{10}$$

where $E[\text{Sales}(K, p)] = E[\text{Demand}(p)] - E[\text{Shortages}(K, p)]$. This representation of the profit function was first introduced by Petruzzi and Dada (1999).

The first-order conditions of (9) and (10) written in this representation are

$$(w - c_k - c_p)\frac{dE[\text{Sales}(K, p)]}{dK} - c_k\frac{dE[\text{Unused inventory}(K, p)]}{dK} = 0$$

$$E[\text{Sales}(K, p)] + (p - w)\frac{dE[\text{Sales}(K, p)]}{dp} = 0,$$

and the optimality conditions of the integrated system are

$$(p - c_k - c_p)\frac{dE[\text{Sales}(K, p)]}{dK} - c_k\frac{dE[\text{Unused inventory}(K, p)]}{dK} = 0$$

$$E[\text{Sales}(K, p)] + (p - c_k - c_p)\frac{dE[\text{Sales}(K, p)]}{dp} - c_k\frac{dE[\text{Unused inventory}(K, p)]}{dp} = 0.$$

This perfectly captures the misaligned incentives in the decentralized system. The expected profit of production is the difference between the total contribution expected from sales and the expected loss from the unused inventory. Marketing, however, only cares about the expected sales because production will be responsible for the excess inventory. Also note that for a given K and p the marginal expected sales with respect to K and p, $dE[\text{Sales}(K, p)]/dK$ and $dE[\text{Sales}(K, p)]/dp$, respectively, are the same as for an integrated system. The marginal, expected unused inventory with respect to K, $dE[\text{Unused inventory}(K, p)]/dK$, is also unchanged. What has changed is the independent marketing's marginal cost and the independent production's marginal profit. What the model captures is consistent with traditional views on the dichotomy between production and marketing functions: Marketing looks at the total value of the order placed, while production cares about other order characteristics, such as the volume and the difficulties of fulfilling the orders (Hill 1993). Because $p \geq w \geq c_k + c_p$, from Lemma 1, Equations (7), and (8) we have the following inequalities

$$p_m^K(K) \geq p(K) \quad \text{and} \quad K_p(p) \leq K(p)$$

where $p(K)$ ($K(p)$) is the optimal price (stock) for a given stock (price) in the integrated system.

The second inequality is because the marginal revenue for production is lower in the decentralized system than in the integrated one. The first inequality is caused by a higher marginal cost for independent marketing, as well as the fact that marketing does not incur the cost of leftover inventory. In general, misaligned incentives in decentralized systems are caused by a transfer price and a misallocation of cost. The transfer price distorts the marginal revenue or cost facing the independent decision makers. The result of cost misallocation is that one is responsible for the consequences of decisions made by the other. Jeuland and Shugan (1983) is an example of the former, and a simple two-part tariff coordinates their system. Lee and Whang (1999) discuss the latter: Only the lower echelon party incurs backorder costs, and upstream installations consequently will stock too little in a decentralized system. A simple ''two-part tariff'' contract fails here. In those models where the manufacturing operations are not considered, a decentralized system typically has a higher retail price (Jeuland and Shugan 1983, Spengler 1950) or a lower inventory (Cachon 1998). These results are solely caused by double marginalization in the lower installation, and there is no misallocation of cost involved. In our model, these results do not hold anymore. The comparisons between the decentralized system and the integrated system on inventory level and price are context specific. Take price for example. Although we know that for a same level of inventory, price in the decentralized system is higher, the two systems do not set the same level of inventory. The matter is complicated further as neither $p_m^K(K)$ nor $p(K)$ is in general a monotone function of K.

5. Decentralized Systems: Service Level Games

In practice a commitment from a production department (or an independent manufacturer) can take the form of a level of inventory (e.g., 1,000 units) or a service level (e.g., 90%), with or without some kind of flexibility. It is known from economics that, assuming players use different strategic variables, a game can result in different solutions (Kreps and Scheinkman 1983). Cachon and Zipkin (1999) provide some new developments along this line of thinking in operations management. They show that although there is little *operational* distinction between echelon and local stock policies for a stationary environment in an integrated system, they do differ *strategically.* Although we know that setting an inventory level and a service level in the integrated system are equivalent in the setting we are considering, their strategic distinction has yet to be examined. In this section, we assume that production's strategic variable is the service level z, so the problem facing production is to find a service level to maximize the expected profit for a given retail price while marketing prices to maximize expected profit, assuming that z is fixed. We refer to this as a service level game. That is, production, instead of solving

$$\max_K \Pi_p^K(K),$$

replaces $K - y$ with z and maximizes $\Pi_m^Z(z)$ over z for a fixed p, where

$$\Pi_p^Z(z) = (w - c_k - c_p)\left[(y + \mu) - \int_z^\beta (t - z)\phi(t)\,dt\right]$$
$$- c_k \int_\alpha^z (z - t)\phi(t)\,dt. \tag{11}$$

The reaction function is given by the familiar fractile rule:

$$z_p^Z(p) = \Phi^{-1}\left(\frac{w - c_k - c_p}{w - c_p}\right). \tag{12}$$

Marketing's problem is

$$\max_p \Pi_m^Z(p)$$

where

$$\Pi_m^Z(p) = (p - w)(y + \mu)$$
$$- (p - w)\int_z^\beta (t - z)\phi(t)\,dt. \tag{13}$$

$\Pi_m^Z(p)$ is derived from $\Pi_m^K(p) = (p - w)E\min(K, \tilde{D})$ by substituting $K - y$ by z.

For a given z, marketing will set the retail price as

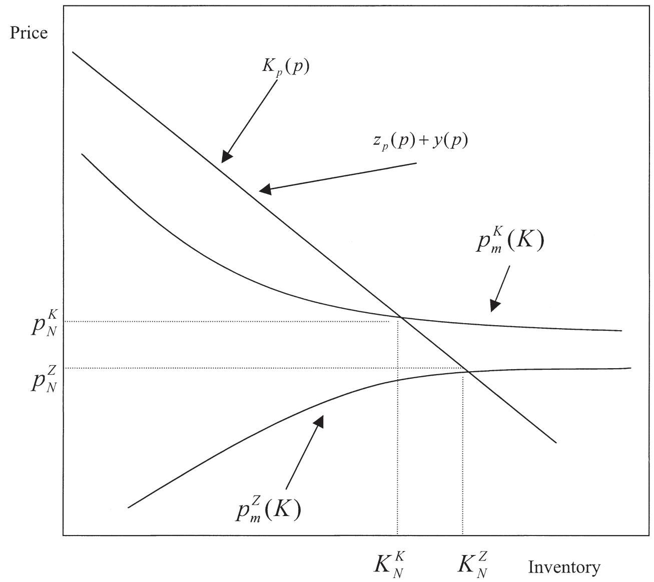
$$p_m^Z(z) = p_m^0 - \frac{\Theta(z)}{2b}. \tag{14}$$

It is easy to show that $p_m^Z(z)$ is increasing and concave in z.

Theorem 2. *(a) There exists a unique Nash solution* $[z_N^Z, p_N^Z]$*, given by (12) and (14). (b)* $z_N^Z = z_N^K$, $p_N^K \geq p_N^Z$*, and* $K_N^K \leq K_N^Z$*, where* $K_N^i = z_N^i + y(p_N^i)$*,* $i = Z, K$.

If we replace z with $K - y(p)$ in (14) and think of $p_m^Z(z)$ as a function of K, we can compare $p_m^Z(z)$ with $p_m^K(K)$ given in (8). It can be easily shown that $p_m^Z(z) \leq p_m^K(K)$; that is, for the same inventory, marketing in the service level game responds with a lower price than that in the inventory game. This should not come as a surprise. Fixing production's strategy in the service level game, marketing can influence the inventory of the system and thus have a greater incentive to lower the retail price. A lower retail price would both increase expected demand and the amount of inventory in the system—i.e., the inventory to serve demand goes up with expected demand. The same is not true, however, for the inventory game. Here a lower retail price merely lowers the payment received on sales that would have been made anyway (because inventory was not expanded) (see Figure 1). A lower incentive to expand demand in the inventory game hurts both production and marketing, as stated in the following theorem. Say *A Pareto dominates B* if everyone prefers *A* to *B* (Bierman and Fernandez 1998), then:

Figure 1 **Nash Equilibria for the Inventory Game and the Service Level Game**



Theorem 3. *The service level game Pareto dominates the inventory game in terms of expected profits.*

Proof. For $i = Z, K$, denote $\Pi_p^i(z_N^i, p_N^i)$ as production's expected profit at equilibrium. Denote $\Pi_m^i(z_N^i, p_N^i)$ for marketing similarly. Because the two games end up with the same optimal z (that is, $z_N^Z = z_N^K$), we have

$$
\begin{aligned}
&\Pi_p^K(z_N^K, p_N^K) \\
&\quad = (w - c_k - c_p)[y(p_N^K) + \mu - \Theta(z_N^K)] - c_k\Lambda(z_N^K) \\
&\quad = (w - c_k - c_p)[y(p_N^K) + \mu - \Theta(z_N^Z)] - c_k\Lambda(z_N^Z) \\
&\quad \le (w - c_k - c_p)[y(p_N^Z) + \mu - \Theta(z_N^Z)] - c_k\Lambda(z_N^Z) \\
&\quad = \Pi_p^Z(z_N^Z, p_N^Z).
\end{aligned}
$$

The inequality is due to $p_N^K \ge p_N^Z$ and because $y(\cdot)$ is a decreasing function. For marketing:

$$
\begin{aligned}
\Pi_m^K(z_N^K, p_N^K) &= (p_N^K - w)[y(p_N^K) + \mu - \Theta(z_N^K)] \\
&= (p_N^K - w)[y(p_N^K) - \mu - \Theta(z_N^Z)] \\
&\le (p_N^Z - w)[y(p_N^Z) + \mu - \Theta(z_N^Z)] \\
&= \Pi_m^Z(z_N^Z, p_N^Z).
\end{aligned}
$$

The inequality follows as p_N^Z maximizes $\Pi_m^Z(z_N^Z, p)$. Finally, because production and marketing are both better off, the firm as a whole is better off if production chooses z rather than K as a strategy. □

To understand this result, note that the key difference between these two games is that marketing in the service level game can induce production to choose a higher inventory level by lowering price. The inventory level is in effect a constraint on marketing's pricing problem. In the service level game, marketing is able to loosen this constraint by reducing price. Put in another way, marketing can price more aggressively without concern about the possible shortage (because expected shortage is determined by z). On the production side, every unit of inventory increased comes with one unit of demand expanded by marketing, so production enjoys more revenue without risking more on unused inventory. As the equilibrium service levels are the same in the two systems, i.e., $z_N^Z = z_N^K$, the two systems result in the same expected shortage and unused inventory. The same level of service is achieved by different means, however. In the service level game the inventory is higher and the expected demand is also higher. In summary, inventory and pricing decisions in a firm have to be coordinated. Inventory will be wasted if demand is weak, while demand will not turn into revenue if inventory is in shortage. The service level game outperforms the inventory game because it allows better coordination of these two decisions. In a sense, it is closer to the integrated system.

This result echoes that of Cachon and Zipkin (1999). The common message is that the operations of a firm with misaligned incentives among functional managers can be strikingly different from those of a centrally controlled firm.

6. Coordinating Contracts

Although the service level game results in improved performance, it does not eradicate the misaligned incentives. A natural question for HQ is how to design a scheme such that the functional objectives are aligned with the firm's objective. The following theorem demonstrates that such coordination can be achieved.

THEOREM 4. *HQ can coordinate the replenishment and pricing in both games by a mechanism of the following form: (a) a transfer price per unit is set to be $w(y) = (1 - \gamma)p(y) + \gamma(c_k + c_p)$, where $p(y)$ is the inverse demand function and γ is an arbitrary number satisfying $\gamma \in [0, 1]$ and (b) marketing pays a portion γ of the cost of unused inventory.*[3]

PROOF. To show that the scheme is a coordinating contract, we need to show $[p^I, z^I]$ is a Nash equilibrium for the decentralized system. That is, for the service level game, we need to show that the following two equations hold:

$$
z^I = \operatorname*{argmax}_z \Pi_p^Z(p^I, z) \quad \text{and}
$$

$$
p^I = \operatorname*{argmax}_z \Pi_m^Z(p, z^I).
$$

With the proposed transfer payment scheme, production's expected profit is:

[3]Note that the parameter γ determines the allocation of the expected profit between production and marketing, but the choice of γ is irrelevant for coordination. Contracts that coordinate a supply chain but arbitrarily divide the profit among players are common, for example, buy back contract (Pasternack 1985), quantity flexibility contract (Tsay 1999), and revenue sharing (Cachon and Lariviere 2000). See the last paper for a more comprehensive discussion of this issue.

$$\begin{aligned}\Pi_p^Z(p, z) &= [w(y) - c_k - c_p]E[\text{Sales}(z, p)] \\ &\quad - (1-\gamma)c_k E[\text{Unused inventory}(z)] \\ &= [(1-\gamma)p + \gamma(c_k + c_p) - c_k - c_p] \\ &\quad \times E[\text{Sales}(z, p)] \\ &\quad - (1-\gamma)c_k E[\text{Unused inventory}(z)] \\ &= (1-\gamma)\Pi^I(p, z).\end{aligned}$$

And, marketing's expected profit is:

$$\begin{aligned}\Pi_m^Z(p, z) &= [p - w(y)]E[\text{Sales}(z, p)] \\ &\quad - \gamma c_k E[\text{Unused inventory}(z)] \\ &= [p - (1-\gamma)p - \gamma(c_k + c_p)]E[\text{Sales}(z, p)] \\ &\quad - \gamma c_k E[\text{Unused inventory}(z)] \\ &= \gamma\Pi^I(p, z).\end{aligned}$$

So this transfer payment scheme in effect works as a profit-sharing mechanism. As Π_p^Z, Π_m^Z, and Π^I are all linearly related, the conditions for the maximization of Π_p^Z, Π_m^Z, and Π^I are all compatible, i.e., both Π_p^Z and Π_m^Z are concave in their respective decision variables, and the optimal solution (Nash equilibrium) is given by Lemma 1. So $[p^I, z^I]$ is a Nash equilibrium. The proof for the inventory game is similar. □

The contract has two parts: a quantity discount and the sharing of unused inventory. As we know, a quantity discount is a subtle mechanism by which production gives some of the additional profit from additional units to marketing so that marketing has the incentive to sell more, or lower the retail price (see, for example, Jeuland and Shugan 1983). On the other hand, the sharing of the unused inventory is effectively a variant of a buy back contract or return guarantee, which is employed to ensure that production builds the right amount of inventory (see, for example, Emmons and Gilbert 1998, Padmanabhan and Png 1995, and Pasternack 1985). The two parts of this contract eradicate the incentive conflicts caused by double marginalization in marketing and production, respectively. Meanwhile, the role of sharing unused inventory is twofold: It serves not only to cope with the double marginalization problem on production's side, but it effectively makes marketing's incentive with respect to demand uncertainty compatible to the overall firm as well. In other words, being responsible for part of the cost of the unused inventory, marketing finds it in its own interest to lower the price to reduce the expected leftover, just as it did in the integrated system. The second part of the scheme is similar to the ''induced penalty cost'' in Lee and Whang (1999), which is a means to relocate costs in the systems to induce optimal behavior.

Operations management is usually viewed as the management of all the value-added activities in the transformation from input to output. To address the operations issues, one needs to model each player in the process as a value-added entity. The challenge for coordination arises from the transfer price among players and misallocation of costs that affect each player's marginal incentive to perform. Cachon and Zipkin (1999) and Caldentey and Wein (2000) are among the limited number of researchers in supply chain coordination that capture the operations of the higher echelon party.[4] Our observation is that in these more realistic situations, coordination requires more complicated schemes. We have achieved our goal of designing such a scheme for our problem, but the issue of how to trade off the difficulty in implemention and the resulting performance still remains.

7. Numerical Studies

Although there exist contracts such that the integrated solution $[p^I, z^I]$ is a Nash equilibrium, given the complication of implemention, linear transfer pricing is still of interest because of its simplicity in implementation. The rule that stipulates that production commit to a service level instead of an inventory level is especially attractive, not only because it can improve the firm's performance, but also because it improves the performance of both functions so there should be less resistance from functional managers

[4]If we think of both the quantity discount and sharing of unused inventory as transfer payments between production and marketing, the scheme here appears to be an application of that proposed in Caldentey and Wein (2000).

Table 1 **Comparing the Three Systems with Different Transfer or Price** (w)

w	Π^I	Π^Z	$(\Pi^I-\Pi^Z)/\Pi^Z$	Π^K	$(\Pi^Z-\Pi^K)/\Pi^K$
4.00	526.10	508.96	3.37%	490.29	3.81%
5.00	526.10	511.80	2.79%	498.66	2.64%
6.00	526.10	508.32	3.50%	497.13	2.25%
7.00	526.10	501.04	5.00%	490.94	2.06%
8.00	526.10	490.78	7.20%	481.47	1.93%

Table 2 **Comparing the Three Systems with Different Demand Sensitivity** (b)

b	Π^I	Π^Z	$(\Pi^I-\Pi^Z)/\Pi^Z$	Π^K	$(\Pi^Z-\Pi^K)/\Pi^K$
2.50	1012.81	965.54	4.90%	901.16	7.14%
3.00	814.40	777.85	4.70%	723.33	7.54%
3.50	673.30	643.99	4.55%	596.64	7.94%
4.00	567.99	543.79	4.45%	501.90	8.35%
4.50	486.51	466.02	4.40%	428.46	8.77%

during implementation.[5] A natural question then is how close to the integrated solution is the performance of decentralized systems if HQ uses linear transfer pricing. Alternatively, under what circumstances should more sophisticated contracts be considered? We have undertaken extensive numerical studies to measure the improvement in total expected profits when changing from the inventory game to the service level game, and the differences in the expected profits between the integrated and decentralized systems. The conclusions from these studies are context and data specific.

Table 1 shows the expected profits of the three systems when the transfer price w is varied. In this experiment, $y(p) = 100 - 4.5p$, $\tilde{\epsilon}$ is exponentially distributed with rate 0.1, $c_k = 0.5$, and $c_p = 2$. The total expected profits of the service level game and inventory game increase initially and then decrease in w. This behavior is jointly caused by the following two effects. On the one hand, when w increases, the marginal cost marketing sees is further distorted from the that of firm (i.e., c_k). This would drive the retail price, which is already too high, even higher. On the other hand, when w increases, production has a higher critical fractile, which leads to a higher service level. The first effect enlarges the gaps between different systems while the second reduces them. In Table 2, $w = 5$, $c_k = 1.1$, and other parameters are the same as those in Table 1, except that the demand sensitivity is varied.

We compare the three systems when the standard deviation (or coefficient of variation) of the demand

[5]In Cachon and Zipkin (1999), production prefers local inventory tracking, but marketing's preference depends on the parameters of the game.

Table 3 **Comparing the Three Systems with Different** CV (β)

β	Π^I	Π^Z	$(\Pi^I-\Pi^Z)/\Pi^Z$	Π^K	$(\Pi^Z-\Pi^K)/\Pi^K$
1.00	567.99	543.79	4.45%	501.90	8.35%
1.20	561.81	525.46	6.92%	484.78	8.39%
1.40	554.33	503.97	9.99%	464.71	8.45%
1.60	545.68	479.61	13.78%	441.95	8.52%
1.80	535.90	452.65	18.39%	416.77	8.61%
2.00	525.02	423.38	24.01%	389.44	8.72%

randomness is varied in Table 3. In these experiments, $y(p) = 100 - 4p$, $\tilde{\epsilon}$ has the following distribution:

$$\Phi(t) = \begin{cases} 1 - \exp\left\{-\dfrac{1}{\mu}\left[\dfrac{t-(1-\beta)\mu}{\beta}\right]\right\} & \text{if } t \geq (1-\beta)\mu \\ 0 & \text{if } t < (1-\beta)\mu. \end{cases}$$

Note that $\tilde{\epsilon}$ is generated by applying a mean preserving transformation to a random variable with negative exponential distribution. Here β can take an arbitrary nonnegative value. When the parameter β varies, the mean is unchanged but the variance (and coefficient of variation) is changed. We take different values for β and compare the three systems. Other parameters are chosen as: $\mu = 10$, $w = 5$, $c_p = 2$, and $c_k = 1.1$. Within this context, the differences in expected profits both between the integrated system and the service level game and between the two decentralized systems increase with β. The magnitude of the increases is not the same, however. When β increases from one to two, the gap between the integrated system and the service level game increases from 4.45 to about 24%, while the gap between the service level game and the inventory game increases from 8.35 to only 8.72%. The data suggest that the higher the demand variance is, the wider the gaps between different systems are. This can be explained by the fact that the decentralized systems have more room for improvement when the demand is more variable.

8. Strategic Dominance

Production and marketing functions are often given different power in firms. In a marketing dominant firm, the role of production is to react to the demand placed upon it: Marketing announces the price first, then production prepares the inventory afterward, knowing the pricing decision from marketing. After the demand is realized, production fulfills the orders in its own best interests. In a production dominant firm, however, production decides how many units to stock (or produce) first, anticipating the reaction from marketing. Then marketing chooses a price to sell the inventory. These situations are best analyzed by Stackelberg games. We seek subgame perfect equilibria, i.e., the second player chooses an optimal response to the first player's strategy, and the first player (correctly) anticipates this behavior.[6]

Production Stackelberg

This is the situation where production is the dominant function. Production announces a choice of service level first, then marketing will decide on price according to production's choice on service level; that is, marketing will choose a price given by (14). By treating p as a function of z, production will still maximize (11). The concavity of $\Pi_p^Z(p_m^Z(z), z)$ can be verified, so the best z, denoted z_{PS}^Z, for production is determined by the first-order condition:

$$\Phi(z_{PS}^Z) = (w - c_k - c_p)/(w - c_p + c_k), \qquad (15)$$

and marketing will choose a price by

$$p_{PS}^Z = p_m^0 - \Theta(z_{PS}^Z)/(2b).$$

THEOREM 5. *(a)* $z_{PS}^Z \le z_N^Z$. *(b)* $p_{PS}^Z \le p_{PN}^Z$. *(c)* $K_{PS}^Z \le K_N^Z$.

[6]The sequential game should be distinguished from the postponement situations described in Van Mieghem and Dada (1999), where decisions are made with different knowledge of the uncertain demand. Here both stocking and pricing are made with the common forecast to the demand, but the follower makes his/her decision after observing the leader's decision.

In finding an optimal service level, production trades off the cost of underage and overage. Compared to the Nash games, a dominant production has a lower marginal incentive to provide a service level because the service level negatively influences the expected demand via price increases by marketing because $p_m^Z(z)$ is increasing, while in Nash games the expected demand is assumed to be fixed (because the price is fixed) when searching for the best service level. (For a similar reason, the price is lower when production moves first in the inventory game). Therefore, production Stackelberg games provide a lower service level. Production is better off having higher expected demand while providing less inventory, and marketing is worse off by being provided with a lower service level. The firm as a whole is worse off. Formally,

THEOREM 6. *(a)* $\Pi_p^Z(z_{PS}^Z,\ p_{PS}^Z) \ge \Pi_p^Z(z_N^Z,\ p_N^Z)$. *(b)* $\Pi_m^Z(z_{PS}^Z, p_{PS}^Z) \le \Pi_m^Z(z_N^Z, p_N^Z)$. *(c)* $\Pi^Z(z_{PS}^Z, p_{PS}^Z) \le \Pi^Z(z_N^Z, p_N^Z)$.

The first part is from the fact that z_N^Z is a feasible solution in production's expected profit maximization problem in the production Stackelberg game, so production enjoys a first-mover advantage. For marketing,

$$\Pi_m^Z(z_{PS}^Z, p_{PS}^Z) = \max_p \Pi_m^Z(z_{PS}^Z, p) \le \max_p \Pi_m^Z(z_N^Z, p)$$
$$= \Pi_m^Z(z_N^Z, p_N^Z).$$

The inequality is because Π_m^Z is increasing in service level for a given price and $z_{PS}^Z \le z_N^Z$. The last part demonstrates that the firm as a whole would prefer the two functions to have equal influence rather than a dominant production function, as the dominant production would induce marketing to generate more demand while reducing the inventory level. The degraded performance is caused by this suboptimal behavior from production. The proof for this is more involved and is provided in the Appendix.

Marketing Stackelberg

In the service level game, as the reaction function (12) is independent of p; it is easy to show that the Nash solution is the unique marketing Stackelberg equilibrium. So from the last part of Theorem 6 we have $\Pi^Z(z_{PS}^Z, p_{PS}^Z) \le \Pi^Z z_N^Z, p_N^Z) = \Pi^Z(z_{MS}^Z, p_{MS}^Z)$. In the inventory game, note that for an announced retail price production chooses an inventory according to (7). Anticipating that production is going to behave in such a way, marketing sets a price to maximize $\Pi_m^K(p) = (p-w)E[\text{Sales}(p, K)]$, subject to production's incentive compatibility constraint. The following result follows easily by comparing the solution of this constrained optimization problem with the corresponding Nash solution.

THEOREM 7. *(a)* $z_{MS}^K = z_N^Z$. *(b)* $p_{MS}^K = p_N^Z$.

In the service level game, marketing can influence production's decision in setting inventory because production has to adjust the inventory level when the price changes to maintain a given service level. Similarly, in a marketing Stackelberg game, marketing can influence production by announcing the pricing decision. This is the rationale behind the equivalence of these two games. As the service level game Pareto dominates the inventory game when the two players simultaneously choose their decisions, this theorem is equivalent to the following statement: In the inventory game, having a dominant marketing function will benefit both production and marketing and, thus, the firm as a whole. This has interesting managerial implications. When production is committed to a service level (i.e., in the service level game), there is no operational distinction whether marketing is the leader of the supply chain or the two functions have equal power. However, when production is committed to a level of inventory (i.e., in the inventory game), marketing should be the leader. As the lead marketing can induce production to build a higher level of inventory by announcing the same retail price as that in the service level game. This result serves as a concrete proposal to achieving the outcomes of the service level Nash game: an organizational change so that marketing becomes the leader.

The purpose of analyzing Stackelberg games in current research is to understand the internal operations when a firm does have a dominant function. The questions of whether there should be a dominant function and who should be the dominant function are clearly more important. We do not intend to give a simple and general answer here, neither can the model provide one. The recommendation that marketing should be the dominant function should be applied with care because there are so many factors that might affect the choice of a dominant function (if at all), especially the balance of demand and capacity in the industry (Hill 1993). Also, a richer model of production operations, such as including a choice of effort level to reduce costs or improve quality, may endow production with more opportunity to influence the performance. Despite this caveat, the results from the Stackelberg games at least suggest functional dominance is indeed relevant to the firm's operations. In particular, having a dominant function tends to end up with an equilibrium price no higher than the Nash counterpart.[7] In addition, having a dominant function can lead to a different service level and inventory. Internal schemes (e.g., the choice between a linear transfer price or a more sophisticated mechanism, the choice between a commitment on inventory or service level) that are used for setting performance measures for functional managers play a significant role in determining whether there should be a dominant function and who it should be.

9. Conclusions and Suggestions for Future Research

There are many potential advantages to delegating decision making to functional managers. However, unless care is taken to design the right mechanism to motivate the resulting independent entities, much of the efficiency possible through joint coordination will be lost. The measures needed to coordinate cross functions (that is, to ensure incentives are in place to mitigate this loss) appear to depend strongly on exactly what the various parts of the supply chain actually do, what value they add. This paper considers two particular fundamental functions of firms: inventory replenishment by production and retail pricing by marketing in advance of demand information. This paper provides a simple framework in which to understand the conflicts and makes it explicit that uncertainty hurts both marketing and production, but

[7] That is, $p_{PS}^Z \le p_N^Z$, $p_{PS}^K \le p_N^K$, $p_{MS}^K = p_N^Z \le p_N^K$, and $p_{MS}^Z = p_N^Z$.

in very different ways. A shortage cost and overage cost are both incurred by production, but marketing only incurs a shortage cost. The coordination contract proposed in this paper has a practical interpretation: Production passes on part of the cost of the unused inventory and provides a quantity discount to marketing so that both production's and marketing's incentive are compatible to that of the firm. This type of contract in one form or another is not uncommon in practice.

Our aim is to provide insights into a setting where the downstream player's marketing activities affect the demand uncertainty in a well-defined way, while the upstream player's cost is affected by the uncertainty in a way different from that of the downstream player. Other than the additive demand uncertainty, a multiplicative form (that is, $D(p, \tilde{\epsilon}) = y(p)\tilde{\epsilon}$) is also common in the literature. The main distinction between these two representations of uncertainty is that the variance increases but the coefficient of variation is unchanged when the price decreases in the multiplicative case, while in the additive case the coefficient of variation decreases but the variance is unchanged. These behaviors are not uncommon in practice. The representation of uncertainty may sometimes be critical. For example, how the introduction of uncertainty into the classical monopoly pricing model affects the optimal price depends on the nature of the uncertainty (Karlin and Carr 1962). To achieve robustness in the results, we have done a parallel analysis assuming a multiplicative demand; the results are available in Li and Atkins (2001). In this case, a service level with a similar interpretation is defined as: $z = K/y(p)$. It turns out that these two models do provide consistent qualitative results, in particular, in the comparison between the service level game and inventory game, the comparison between the Nash game and the marketing Stackelberg game, and the coordinating contract.

We have assumed a specific functional form for the deterministic demand function for ease of analysis. Other demand functions in the literature, for example, $y(p) = ap^{-b}$ or $y(p) = ae^{-bp}$, for the multiplicative case have been experimented. Other than the conditions for guaranteeing the uniqueness of the integrated systems being slightly different, the qualitative results are very much the same. A big step forward would be to consider a demand function that has both additive and multiplicative uncertainty. Before making this extension, there are two issues that need to be addressed. First, it would need to be established that a more general demand function is needed for real-world modeling and what additional characteristics can only be captured by such a model. Secondly, can the structural results of the optimal policy yield managerial insights in the integrated system?

In our model the value added by production is reflected through the replenishment decision. There are many alternatives to this choice. One may be providing after-sale services or deciding on whether and to what extent product differentiation should be delayed, etc. The marketing function can also be extended to include other marketing decisions. Extensions within the same setting would be made by considering multiple products that are jointly managed, or multiple periods so that the dynamic nature of operations is captured. We believe that the multiple period versions of this problem are potentially more rich. Based on our observation of the existing research concerning the dynamic version of the newsvendor problem with price effects, extending our model to a dynamic setting should not be an easy task. Just as Petruzzi and Dada (1999) point out, ''. . . In stark contrast to the research on the single period model, the literature on multiple period models does not provide structural results of the optimal policy that yields managerial insight; rather the emphasis has been on technical properties that speed up computation of optimal policies'' (p. 192). The major challenge for incorporating dynamics into our model is likely to be: how to make assumptions that can ease the computation of optimal policies without sacrificing managerial insight.

Appendix

Proof of Lemma 1. The proof involves two steps.

Step 1. (p^I, K^I) satisfies the first-order conditions of (2).

$$\left.\frac{\partial G^I(p, K)}{\partial K}\right|_{p=p^I, K=K^I} = (p^I - c_p\bar{\Phi}[K^I - y(p^I)] - c_k$$

$$= 0$$

and

$$\left.\frac{\partial G^I(p, K)}{\partial p}\right|_{p=p^I, K=K^I} = -2b(p^I - p^0) - \Theta[K^I - y(p^I)] - bc_k$$

$$+ b(p^I - c_p)\bar{\Phi}[K^I - y(p^I)]$$

$$= -b\{c_k - (p^I - c_p)\bar{\Phi}[K^I - y(p^I)]\}$$

$$= 0.$$

Step 2. We show the optimal solution of the program (2) is indeed unique and given by its first-order conditions. Because $G^I(p, K)$ is concave in each of its two variables, we can solve the problem sequentially. Suppose we first find the best K for a given p and denote it $K(p)$, then

$$K(p) = y(p) + \bar{\Phi}^{-1}\left(\frac{c_k}{p - c_p}\right).$$

Treating $G^I[p, K(p)]$ as a function of p and taking derivatives:

$$\frac{dG^I[p, K(p)]}{dp} = -2b(p - p^0) - \Theta[K(p) - y] - bc_k$$

$$+ b(p - c_p)\bar{\Phi}[K(p) - y]$$

$$= -2b(p - p^0) - \Theta[K(p) - y].$$

Define $R(p) = -2b(p - p^0) - \Theta[K(p) - y(p)]$, then

$$R'(p) = -2b + \bar{\Phi}[K(p) - y]\left(\frac{dK(p)}{dp} + b\right)$$

$$= -2b + \frac{\bar{\Phi}[K(p) - y]}{(p - c_p)h[K(p) - y]}$$

and

$$R''(p) = -\left\{\frac{\bar{\Phi}[K(p) - y]}{(p - c_p)^2 h[K(p - y]}\right.$$

$$+ \frac{\bar{\Phi}[K(p) - y]h'[K(p) - y]}{(p - c_p)h^2[K(p) - y]}\left(\frac{dK(p)}{dp} + b\right)$$

$$\left. + \frac{\phi[K(p) - y]}{(p - c_p)h[K(p) - y]}\left(\frac{dK(p)}{dp} + b\right)\right\}$$

$$\leq 0.$$

The inequality follows as $h'(\cdot)$ is nonnegative and

$$\frac{dK(p)}{dp} + b \geq 0.$$

So $R(p)$ is unimodal. And

$$R(p)|_{p\to c_k+c_p} = 2b(p^0 - c_k - c_p) - \Theta(\alpha)$$

$$= a - b(c_k + c_p) - \mu + \alpha = y(c_k + c_p) + \alpha$$

$$> 0.$$

The inequality follows from the assumption of nonnegativity of demand, and

$$R(p)|_{p\to\infty} = -\infty.$$

So $R(p)$ has only one root and $G^I[p, K(p)]$ is unimodal, first increasing and then decreasing. Therefore, given that $\tilde{\epsilon}$ has an increasing failure rate, the program (2) has a unique solution given by its first-order condition. □

Proof of Lemma 2. It is easy to see that (p^I, z^I) satisfies the first-order conditions of (4). We only need to show that the optimal solution of (4) is indeed given by its first-order conditions. From the chain rule

$$\frac{d\Pi[p(z), z]}{dz} = p^0 - \frac{\Theta(z)}{2b} - c_k - c_p - \left(p^0 - \frac{\Theta(z)}{2b} - c_p\right)\Phi(z).$$

To identify values of z that satisfy this first-order optimality condition, let $R(z) = d\Pi[p(z), z]/dz$ and consider finding the zeros of $R(z)$:

$$\frac{dR(z)}{dz} = \bar{\Phi}(z)\left[\frac{\bar{\Phi}(z)}{2b} - \left(p^0 - \frac{\Phi(z)}{2b} - c_p\right)h(z)\right]$$

where $h(z)$ is the hazard rate. If $dR(z)/dz = 0$ has no root, then $R(z)$ is monotone; otherwise,

$$\left.\frac{d^2R(z)}{dz^2}\right|_{[dR(x)/dz]=0} = -\left(p^0 - \frac{\Theta(z)}{2b} - c_p\right)h'(z)\bar{\Phi}(z)$$

$$= -\frac{\bar{\Phi}^2(z)h'(z)}{2bh(z)} \leq 0$$

for all $\tilde{\epsilon}$ with IFR. So $R(z)$ is either monotone or unimodal, implying that $R(z)$ has at most two roots. Further, $R(\beta) = -c_k \leq 0$. Therefore, if $R(z)$ has only one root, it indicates a change of sign for $R(z)$ from positive to negative, and thus it corresponds to a local maximum of $\Pi[p(z), z]$. To ensure $R(z)$ has only one root, a sufficient condition is $R(\alpha) \geq 0$; that is

$$R(\alpha) = p^0 - \frac{\Theta(\alpha)}{2b} - c_k - c_p = p^0 - c_k - c_p - \frac{\mu - \alpha}{2b}$$

$$= \frac{1}{2b}[y(c_k + c_p) + \alpha] \geq 0.$$

The inequality holds as we assumed $y(c_k + c_p) + \alpha \geq 0$ to avoid negative demand. So the optimal solution is given by the first-order condition and the result follows. □

Proof of Theorem 2. The reaction function $z_p^Z(p)$ is a constant, and $p_m^Z(z)$ is increasing and concave in z with upper bound p_m^0 and lower bound $p_m^0 - \mu + \alpha$. It follows that the two reaction functions cross once and only once. Existence and uniqueness are thus established. The second part of the theorem can be easily verified by comparing the optimal conditions of the games. □

Proof of Theorem 5. By (12) and (15), we have $z_{PS}^Z \leq Z_N^Z$, and $p_m^Z(z) = p_m^0 - \Theta(z)/(2b)$ is increasing in z, so $p_{PS}^Z = p_m^Z(z_{PS}^Z) \leq p_m^Z(z_N^Z) = p_N^Z$. Write K as a function of z. That is, $K(z) = z + y(p_m^Z(z)) = z + a - bp_m^0 + \Theta(z)/2$. $K(z)$ is increasing on z, so $K_{PS}^Z \leq K_N^Z$ follows as $z_{MS} \leq z_N^Z$. □

Proof of Theorem 6 (Part 3). As marketing has the same reaction functions regardless of Nash game or production Stackelberg,

$\Pi^Z(z_{PS}^Z, p_{PS}^Z)$ and $\Pi^Z(z_N^Z, p_N^Z)$ can be expressed as the values of a common function evaluated at different points. That is if we define

$$\Pi(z) = \left[p_m^0 - \frac{\Theta(z)}{2b} - c_k - c_p\right]\left[a + \mu - bp_m^0 - \frac{\Theta(z)}{2}\right] - c_k\Lambda(z),$$

then $\Pi^Z(z_{PS}^Z, p_{PS}^Z) = \Pi(z_{PS}^Z)$ and $\Pi^Z(z_N^Z, p_N^Z) = \Pi(z_N^Z)$. We proceed with the proof in two steps.

(a) $\Pi(z)$ is unimodal, first increasing then decreasing:

$$\frac{d\Pi(z)}{dz} = \frac{\bar{\Phi}(z)}{2b}[a + \mu - \Theta(z) - b(c_k + c_p)] - c_k\Phi(z).$$

Let $R(z) = d\Pi(z)/dz$, then

$$R'(z) = \frac{\bar{\Phi}(z)}{2b}\{\bar{\Phi}(z) - h(z)[a + \mu - \Theta(z) - b(c_k + c_p)]\}$$

$$R''(z)|_{R'(z)=0} = -\phi(z) - h(z)\bar{\Phi}(z) - h'(z)[a + \mu - \Theta(z) - b(c_k + c_p)] \le 0,$$

so $R(z)$ is either unimodal or monotone, implying that $R(z)$ has at most two roots. Further, $R(\alpha) = (1/2b)[a - b(c_k + c_p) + \alpha] \ge 0$, $R(\beta) = -c_k \le 0$. This indicates that $R(z)$ has only one root and that it changes its sign from positive to negative. Thus it corresponds to a global maximum of $\Pi(z)$. Let $R(z)|_{z=z^*} = 0$.

(b) We show $z_{PS}^Z \le z^*$ and $z_N^Z \le z^*$.

$$R(z)|_{z=z_{PS}^Z} = \frac{c_k}{b(w - c_p + c_k)}[a + \mu - \Theta(z_{PS}^Z) - bw] = \frac{2c_k}{(w - c_p + c_k)}(p_{PS}^Z - w) \ge 0$$

and

$$R(z)|_{z=z_N^Z} = \frac{c_k}{2b(w - c_p)}[a + \mu - \Theta(z_N^Z) - 2bw + p(c_p + c_k)] = \frac{c_k}{(w - c_p)}(p_N^Z - w) \ge 0,$$

so $z_{PS}^Z \le z^*$ and $z_N^Z \le z^*$. Furthermore, because $z_{PS}^Z \le z_N^Z$ (Theorem 5), we have $\Pi^Z(z_{PS}^Z, p_{PS}^Z) = \Pi(z_{PS}^Z) \le \Pi^Z(z_N^Z, p_N^Z) = \Pi(z_N^Z)$. □

References

Amihud, Y., H. Mendelson. 1983. Price smoothing and inventory. *Rev. Econom. Stud.* **50** 87–98.

Anupindi, R., Y. Bassok. 1998. Supply contracts with quantity commitments and stochastic demand. S. Tayur, M. Magazine, and R. Ganeshan, eds. *Quantitative Models for Supply Chain Management.* Kluwer Academic Publishers, Boston, MA.

Bierman, H., L. F. Fernandez. 1998. *Game Theory with Economic Applications.* Addison Wesley, Reading, MA.

Cachon, G. 1998. Competitive supply chain inventory management. S. Tayur, M. Magazine, and R. Ganeshan, eds. *Quantitative Models for Supply Chain Management.* Kluwer Academic Publishers, Boston, MA.

——, M. A. Lariviere. 2000. Supply chain coordination with revenue sharing contracts: Strengths and limitations. Working paper, Duke University, North Carolina.

——, P. H. Zipkin. 1999. Competitive an cooperative inventory policies in a two-stage supply chain. *Management Sci.* **45** 936–953.

Caldentey, R., L. Wein. 2000. Analysis of a production inventory system. Working paper, Massachusetts Institute of Technology, Cambridge, MA.

Chen, F. 2000a. Sales incentives, market information, and production/inventory planning. Working paper, Graduate School of Business, Columbia University, New York.

——. 2000b. Market segmentation, advanced demand information, and supply chain performance. Working paper, Graduate School of Business, Columbia University, New York.

Clark, A., H. Scarf. 1960. Optimal policies for a multi-echelon inventory problem. *Management Sci.* **6** 475–490.

De Groote, X. 1994. Flexibility and the marketing/manufacturing interface. *Internat. J. Production Econom.* **36** 153–167.

Eliashberg, J., R. Steinberg. 1993. Marketing-production joint decision making. J. Eliashberg and G. L. Lilien, eds. *Marketing: Handbook in OR and MS.* North-Holland, Amsterdam, The Netherlands.

Emmons, H., S. M. Gilbert. 1998. The role of returns policies in pricing and inventory decisions for catalogue goods. *Management Sci.* **44** 276–283.

Ernst, R. L. 1970. A linear inventory model of a monopolistic firm. Ph.D. dissertation, Dept. of Economics, University of California, Berkeley, CA.

Federgruen, A., A. Heching. 1999. Combined pricing and inventory control. *Oper. Res.* **47** 454–475.

Hill, T. 1993. *Manufacturing Strategy.* MacMillan, Basingstoke, England.

Jeuland, A., S. M. Shugan. 1983. Managing channel profits. *Marketing Sci.* **2** 239–272.

Karlin, S., C. R. Carr. 1962. Prices and optimal inventory policy. K. Arrow, S. Karlin, and H. Scarf, eds. *Studies in Applied Probability Theory and Management Science.* Stanford University Press, Stanford, CA.

Kreps, D., and J. Scheinkman. 1983. Quantity precommitment and Bertrand competition yield Cournot outcomes. *Bell J. Econom.* **14** 326–337.

Lee, H. L., S. Whang. 1999. Decentralized multi-echelon supply chains: Incentive and information. *Management Sci.* **45** 633–640.

Li, Q., D. Atkins. 2001. Coordinating replenishment and pricing in a firm. Working paper, The University of British Columbia, Vancouver, Canada.

Nahmias, S. 1989. *Production and Operations Analysis.* Richard D. Irwin, Boston, MA.

Padmanabhan, V., I. Png. 1995. Returns policies: Make money by making good. *Sloan Management Rev.* **37** (Fall) 65–72.

Pasternack, B. 1985. Optimal pricing and returns policies for perishable commodities. *Marketing Sci.* **4** 166–176.

Petruzzi, N. C., M. Dada. 1999. Pricing and the newsvendor problem: A review with extensions. *Oper. Res.* **47** 183–194.

Porteus, E. L. 1990. Stochastic inventory theory. D. P. Heyman and

M. J. Sobel, eds. *Stochastic Models.* North-Holland, Amsterdam, The Netherlands.

——, S. Whang. 1991. On manufacturing/marketing incentives. *Management Sci.* **37** 1166–1181.

Spengler, J. 1950. Vertical integration and antitrust policy. *J. Political Econom.* **58** 347–352.

Stern, L. W., A. I. El-Ansary. 1992. *Marketing Channels.* Prentice Hall, Inc., Upper Saddle River, NJ.

Tsay, A. A. 1999. The quantity flexibility contract and supplier-customer incentives. *Management Sci.* **45** 1339–1358.

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

Whitin, T. M. 1955. Inventory control and price theory. *Management Sci.* **2** 61–68.

The consulting Senior Editor for this manuscript was Evan L. Porteus. This manuscript was received on May 29, 2001, and was with the authors 170 days for 3 revisions. The average review cycle time was 82 days.