Manufacturing & Service Operations Management

Managing Inventory Over a Short Season: Models with Two Procurement Opportunities

Jian Li, Suresh Chand, Maqbool Dada, Shailendra Mehta,

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 11, No. 1, Winter 2009, pp. 174–184
ISSN 1523-4614 | EISSN 1526-5498 | 09 | 1101 | 0174

DOI 10.1287/msom.1070.0204

Managing Inventory Over a Short Season: Models with Two Procurement Opportunities

Jian Li
College of Business and Management, Northeastern Illinois University, Chicago, Illinois 60625, jli@neiu.edu

Suresh Chand, Maqbool Dada
Krannert School of Management, Purdue University, West Lafayette, Indiana 47907
{suresh@purdue.edu, dada@mgmt.purdue.edu}

Shailendra Mehta
Indian Institute of Management, Ahmedabad, Vastrapur 380015, India, shailendra@iimahd.ernet.in

Motivated by many recent applications reported in the literature, we examine the impact of a second procurement opportunity on inventory management of products with short selling seasons. In our framework, the first order is placed at the start of the preseason and delivered at the start of the selling season; the second order is placed at or after the start of the selling season for subsequent delivery. Under this framework, the decision maker must make three interrelated choices: the first order quantity, when to place the second order, and the second order quantity. Our focus is on elucidating the optimal policy structure for the three interrelated decisions. By casting our models as sequential decision-making problems, we are able to reduce the optimization problems into sequential and embedded searches for the concerned decision variables that allow us to identify the conditions on the economic parameters and demand distribution to effectively facilitate the search for the optimal solutions.

Key words: newsvendor problem; short selling season; second order opportunity; inventory management; optimal policy
History: Received: April 28, 2005; accepted: November 6, 2007. Published online in *Articles in Advance* April 25, 2008.

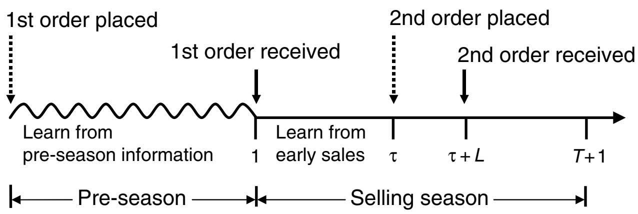
1. Introduction

Traditionally, in many industrial settings ranging from agriculture to fashion, the decision of how much to procure is made well before the start of a well defined selling season. Consequently, the decision maker is unable to take advantage of the subsequent information that becomes available as the season draws closer. As exemplified by its pioneering use in the 1980s by Benetton, the fashion retail giant, providing a second order opportunity around the start of the season can significantly reduce markdowns and leftover inventory. Propelled by this and related innovations in supply chain management, researchers have begun a re-examination of analytical models of inventory management that support such decision making. The range of applications has been extensive. It includes models on quick response (Fisher and Raman 1996), catalog sales (Eppen and Iyer 1997), planning hybrid seed inventories (Jones et al. 2001), and electric power (Iyer et al. 2003).

Reflecting the extensive range of applications, these models with a second order opportunity may be viewed as special cases of the general model depicted in Figure 1, which depicts the season as consisting of T periods. The first order is placed at the start of the preseason and delivered at the start of the selling season; the second order is placed at or after the start of the selling season for subsequent delivery. Under this framework, the decision maker must make three interrelated choices: the first order quantity, when to place the second order, and the second order quantity.

The three decisions to be made are influenced by a variety of factors. First, reflecting the underlying business scenarios, it is appropriate in many settings to

Figure 1



use the preseason to update information on the distribution of demand (Donohue 2000). Second, reflecting other business settings, initial demand information during the season can be used to update the demand distribution. Under these settings, as in Eppen and Iyer (1997), at the time of placing the second order, the decision maker may not only have refined information on demand for the remainder of the season, but also has the benefit of including the current inventory level to further fine tune the order quantity. And, third, these decisions are influenced by how a stockout is handled: should demand be backordered until the second order is received, as in Fisher et al. (2001), or should it only be satisfied from on-hand inventory, as in the models considered by Milner and Kouvelis (2002)?

Our goal in this paper is to conduct a comprehensive analysis of the stocking problem with two order opportunities captured by the framework depicted in Figure 1, with focus on elucidating the optimal policy structure. We begin in §2 with a generalization of the model considered by Fisher et al. (2001). In this model, the first order is placed prior to the start of the season and received by the retailer at the start of period 1 of the selling season. If sufficient inventory has been depleted at the start of period τ, the retailer would place the second order for delivery at the start of period $\tau + L$. All demand that may have been backordered before the start of period τ will be filled by the second order, and demand after the start of period τ is accepted until all inventory on hand and on order is depleted. Our representation of the model generalizes the model of Fisher et al. (2001) because it allows for time-dependent backorder costs and holding costs, and it admits demand distributions other than the normal.

After developing the problem formulation, we discern the problem structure that identifies the conditions on the economic parameters and demand distributions to effectively facilitate the search of the optimal solutions (Theorem 1). Our representation of the problem subsumes the special case where the second order is placed at the start of period 1 for delivery at the start of period $L + 1$. This model, by incorporating a positive leadtime, directly generalizes the model of Donohue (2000). We also show that not only can our model accommodate Bayesian updates of demand during the preseason, as in the work of Donohue (2000), it can also accommodate using information on demand collected in the first $\tau - 1$ periods to update the demand distribution used to determine the second order. And then, in §3, we extend the model by allowing the timing of the second order to be determined dynamically from the evolution of the demand process during the season. We provide conditions under which the optimal policy for the second order is of the (s, S) type.

Our findings are summarized in §4. In this section we also discuss the relative position of our research within the rapidly expanding literature on inventory management of goods with short selling periods. While our work directly generalizes the work of Fisher et al. (2001), it also nicely complements the work of Milner and Kouvelis (2002, 2005). Like those of Milner and Kouvelis (2002, 2005), our models vary the timing of the second order. The key differences are that (1) they do not allow for partial backorders, and (2) they focus on examining the interplay between the value of information and flexibility in their three problem variants, whereas we focus on elucidating the problem structure under rather general conditions.

2. Basic Model

We develop a discrete time model of the selling season that consists of T periods numbered $1, 2, \ldots, T$; demand in each period is random and independent across periods. As in Figure 1, Q_0, the first order, is placed at the start of the preseason for delivery at the beginning of period 1. And, Q_τ, the second order, is placed at the beginning of period τ and received at the beginning of period $\tau + L$. In the event that during the first $\tau - 1$ periods the initial inventory, Q_0, is depleted, we continue to accept all additional demand as backorders for the guaranteed delivery at the beginning of period $\tau + L$; during periods τ through $\tau + L - 1$, demand in excess of the on-hand inventory is accepted to be filled from the second order as long as sufficient stock remains in the pipeline. Demand that cannot be filled from the second order is considered lost or, effectively, satisfied from outside the system. Ending inventory at the end of each period is held over for the next period, but any leftover inventory at the end of the season is disposed at a known unit cost. The decision maker must choose these two order quantities to minimize the total expected cost of meeting the demand. This cost is the sum of the purchase cost of the two orders and the expected costs for inventory holding, backorders, lost sales, and disposal of leftover inventory at the end of period T.

We need the following notation and terms to formulate the problem.

2.1. Decision Variables

τ: The period at the beginning of which the second order is placed.

Q_0: The quantity delivered at the beginning of period 1.

Q_τ: The quantity delivered at the beginning of period $\tau + L$.

2.2. Economic Parameters

c_0: The unit purchase cost for Q_0.

c_τ: The unit purchase cost for Q_τ.

c_d: The unit disposal cost charged at the end of the season.

c_u: The unit lost-sales cost.

h_t: The unit holding cost charged on the ending inventory in period t; $t = 1, \ldots, T$.

b_t: The unit backorder cost charged on the cumulative backorders in period t; $t = 1, \ldots, T$.

2.3. Demand Variables

ξ_t: The random demand in period t with probability density function (PDF) $f_t(\cdot)$, cumulative distribution function (CDF) $F_t(\cdot)$, and complementary CDF $\bar{F}_t(\cdot)$.

ξ_{t_1, t_2}: The cumulative demand in periods t_1 through t_2; $1 \le t_1 \le t_2 \le T$;

$$\xi_{t_1, t_2} = \xi_{t_1} + \xi_{t_1+1} + \cdots + \xi_{t_2} \quad (\text{note, } \xi_{t_1, t_1} = \xi_{t_1}).$$

$f_{t_1, t_2}(\cdot)$, $F_{t_1, t_2}(\cdot)$, $\bar{F}_{t_1, t_2}(\cdot)$: The PDF, CDF, and complementary CDF for ξ_{t_1, t_2}, respectively.

2.4. State Variables and Their Dynamics

x_t: The inventory level at the beginning of period t; $t = 1, \ldots, T$.

y_t: The inventory position (x_t + on order in period t) at the beginning of period t; $t = 1, \ldots, T$. These state variables evolve as follows:

$$x_1 = Q_0, \quad x_t = \begin{cases} x_{t-1} - \xi_{t-1} & t \neq 1, \ \tau + L \\ x_{t-1} - \xi_{t-1} + Q_\tau & t = \tau + L, \end{cases}$$

and

$$y_t = \begin{cases} x_t & t = 1, \ldots, \tau - 1 \\ x_t + Q_\tau & t = \tau \\ y_{t-1} - \xi_{t-1} & t = \tau + 1, \ldots, T. \end{cases}$$

Note that the inventory level x_t can be positive or negative and that $y_t \ge x_t$.

2.5. Performance Measures

$B_t(x_t, y_t \mid \xi_t)$: The realized backorder and holding cost in period t with a beginning inventory level x_t and a beginning inventory position y_t for a given demand ξ_t; $t = \tau, \ldots, \tau + L - 1$;

$$B_t(x_t, y_t \mid \xi_t) = \begin{cases} h_t(x_t - \xi_t) & \text{if } 0 \le \xi_t < x_t \\ b_t(\xi_t - x_t) & \text{if } x_t \le \xi_t < y_t \\ b_t(y_t - x_t) & \text{if } y_t \le \xi_t. \end{cases}$$

$B_t(x_t, y_t)$: The expected backorder and holding cost in period t; $t = \tau, \ldots, \tau + L - 1$;

$$\begin{aligned} B_t(x_t, y_t) &= \int_0^\infty B_t(x_t, y_t \mid \xi_t) f_t(\xi_t)\, d\xi_t \\ &= B_t^y(y_t) - B_t^x(x_t), \end{aligned}$$

where

$$\begin{aligned} B_t^x(x_t) &= -h_t \int_0^{x_t} (x_t - \xi_t) f_t(\xi_t)\, d\xi_t \\ &\quad + b_t \int_0^{x_t} \xi_t f_t(\xi_t)\, d\xi_t + b_t \int_{x_t}^\infty x_t f_t(\xi_t)\, d\xi_t; \quad (1) \\ B_t^y(y_t) &= b_t \int_0^{y_t} \xi_t f_t(\xi_t)\, d\xi_t + b_t \int_{y_t}^\infty y_t f_t(\xi_t)\, d\xi_t. \end{aligned}$$

$G_{\tau+L, T}(y_{\tau+L})$: The expected total cost of inventory holding, lost sales, and disposal from periods $\tau + L$

through T with a beginning inventory level $y_{\tau+L}$. (Note that $y_{\tau+L} = x_{\tau+L}$ due to the delivery of Q_τ.)

$$\begin{aligned} G_{\tau+L,T}(y_{\tau+L}) = {} & \sum_{t=0}^{T-\tau-L} h_{\tau+L+t} \int_0^{y_{\tau+L}} (y_{\tau+L} - \xi_{\tau+L,\,\tau+L+t}) \\ & \cdot f_{\tau+L,\,\tau+L+t}(\xi_{\tau+L,\,\tau+L+t})\, d\xi_{\tau+L,\,\tau+L+t} \\ & + c_u \int_{y_{\tau+L}}^{\infty} (\xi_{\tau+L,\,T} - y_{\tau+L}) \\ & \cdot f_{\tau+L,\,T}(\xi_{\tau+L,\,T})\, d\xi_{\tau+L,\,T} \\ & + c_d \int_0^{y_{\tau+L}} (y_{\tau+L} - \xi_{\tau+L,\,T}) \\ & \cdot f_{\tau+L,\,T}(\xi_{\tau+L,\,T})\, d\xi_{\tau+L,\,T}. \end{aligned}$$

$g_t(x_t, y_t)$: The expected total cost of inventory holding, backorders, lost sales, and disposal from periods t through T with a beginning inventory level x_t and a beginning inventory position y_t; $t = \tau, \ldots, \tau+L-1$;

$$g_t(x_t, y_t) = \begin{cases} B_t(x_t, y_t) + \int_0^\infty g_{t+1}(x_t - \xi_t, y_t - \xi_t) f_t(\xi_t)\, d\xi_t \\ \quad t = \tau, \ldots, \tau+L-2 \\ B_t(x_t, y_t) + \int_0^\infty G_{\tau+L,\,T}(y_t - \xi_t) f_t(\xi_t)\, d\xi_t \\ \quad t = \tau+L-1. \end{cases} \tag{2}$$

$\rho_t(x_t)$: The minimum expected total cost of purchase for the second order, inventory holding, backorders in periods t through T, lost sales, and disposals with a beginning inventory level x_t; $1 \le t \le \tau$. For $1 \le t \le \tau - 1$, we have

$$\begin{aligned} \rho_t(x_t) = {} & b_t \int_{x_t}^\infty (\xi_t - x_t) f_t(\xi_t)\, d\xi_t + h_t \int_0^{x_t} (x_t - \xi_t) f_t(\xi_t)\, d\xi_t \\ & + \int_0^\infty \rho_{t+1}(x_t - \xi_t) f_t(\xi_t)\, d\xi_t. \end{aligned} \tag{3}$$

$TC_I(Q_0)$: The minimum expected total cost of the T-period problem with an initial order Q_0.

We are now ready to develop the model. The decision problem for choosing IP_τ at the beginning of period τ, which essentially determines the second order quantity $Q_\tau = IP_\tau - Q_0$, can be formulated as

$$\begin{aligned} \rho_\tau(x_\tau) = \min_{IP_\tau \ge x_\tau} \{ & H_\tau(x_\tau, IP_\tau) = c_\tau(IP_\tau - x_\tau) \\ & + g_\tau(x_\tau, IP_\tau) \}. \end{aligned} \tag{P1.1}$$

The first term of $H_\tau(x_\tau, IP_\tau)$ represents the purchase cost of the second order, and the second term, as defined in (2), represents the sum of the expected cost of inventory holding, backorder, and lost sales in periods τ through T, and the expected cost of disposal of leftover inventory at the end of the season for a beginning inventory level x_τ and a beginning inventory position IP_τ. Similarly, the decision problem for Q_0 can be formulated as

$$\min_{Q_0 \ge 0} \{ TC_I(Q_0) = c_0 Q_0 + \rho_1(x_1 = Q_0) \}. \tag{P1.2}$$

The first term of $TC_I(Q_0)$ represents the purchase cost of the first order, and the second term, as defined in (3), represents the minimum expected total cost in periods 1 through T with a beginning inventory level Q_0, excluding the purchase cost for Q_0. We let $IP_\tau^*(x_\tau)$ be the optimal solution of (P1.1) and Q_0^* be the optimal solution of (P1.2).

Because the first term in (P1.2) is linear, to understand the structure of problem (P1.2) it is sufficient to focus our analysis on the structure of $\rho_1(x_1)$ in (3), which in turn requires an analysis of $H_\tau(x_\tau, IP_\tau)$ by (P1.1). Because $H_\tau(x_\tau, IP_\tau)$ is determined by the expected holding and backorder costs in periods τ through $\tau+L-1$ via $g_t(x_t, y_t)$ and by $G_{\tau+L,\,T}(y_{\tau+L})$, we begin our analysis by characterizing $g_t(x_t, y_t)$, $G_{\tau+L,\,T}(y_{\tau+L})$, $B_t^x(x_t)$, and $B_t^y(y_t)$. (All proofs of Lemmas and Theorems below are provided in the online appendix.)

Lemma 1.

(1) *$G_{\tau+L,\,T}(y_{\tau+L})$ is convex in $y_{\tau+L}$.*

(2) *For $t = \tau, \ldots, \tau+L-1$, $B_t^x(x_t)$ is concave in x_t, and $B_t^y(y_t)$ is concave increasing in y_t.*

(3) *For $t = \tau, \ldots, \tau+L-1$, $g_t(x_t, y_t)$ is separable in x_t and y_t; moreover, $g_t(x_t, y_t)$ is convex in x_t.*

It is important to notice that $B_t(x_t, y_t)$ from (1) has the special structure that, because it is a linear function of the difference of $B_t^y(y_t)$ and $B_t^x(x_t)$, it is separable in x_t and y_t, and that $B_t(x_t, y_t)$ is a "cost" that is the difference of two concave functions, a fact that significantly complicates the analysis of our problem. However, $G_{\tau+L,\,T}(y_{\tau+L})$ is convex in the argument, which significantly facilitates the analysis.

Now, consider the first order optimality condition of $H_\tau(x_\tau, IP_\tau)$ in (P1.1) with respect to IP_τ:

$$\begin{aligned}\frac{\partial H_\tau(x_\tau, IP_\tau)}{\partial IP_\tau} &= c_\tau + \frac{\partial g_\tau(x_\tau, IP_\tau)}{\partial IP_\tau}\\ &= (c_\tau + c_d) + \sum_{t=\tau}^{\tau+L-1} b_t \bar{F}_{\tau,t}(IP_\tau)\\ &\quad + \sum_{t=\tau+L}^{T} h_t F_{\tau,t}(IP_\tau)\\ &\quad - (c_u + c_d)\bar{F}_{\tau,T}(IP_\tau) = 0. \end{aligned} \tag{4}$$

Because the choice of IP_t represents the last chance to place an order, there is a direct connection of (4) to the corresponding first-order condition for the classical newsvendor problem. This connection is most easily seen by first setting the backorder costs and holding costs to zero. When this is the case, the two summations vanish and (4) is precisely the newsvendor condition with the proviso that the uncertain demand is represented by the demand since the time when the order is placed. Our way of charging holding cost results in a weighted average of service levels, but preserves convexity and, therefore, analytical tractability. However, the way of charging backordering costs may breakdown convexity and unimodality.

To see this, consider the second order derivative, which is as follows:

$$-\sum_{t=\tau}^{\tau+L-1} b_t f_{\tau,t}(IP_\tau) + \sum_{t=\tau+L}^{T} h_t f_{\tau,t}(IP_\tau) + (c_u + c_d) f_{\tau,T}(IP_\tau).$$

As can be seen, the second order derivative is not clearly positive or negative, suggesting that $H_\tau(x_\tau, IP_\tau)$ is, in general, not convex or even unimodal in IP_τ and has multiple local minima, making intractable a full characterization. However, there is sufficient structure to come close to a complete characterization. This is because we know from (4) that $H_\tau(x_\tau, IP_\tau)$ as a function of IP_τ may have multiple local minima and maxima, the order-up-to position IP_τ^* depends on x_τ, and, because x_τ is a state variable, IP_τ^* must be defined for all possible realizations of x_τ. As shown in the online appendix, this can be done by iteratively defining IP_τ^i and $a^i (i = 0, 1, \ldots, n)$ based on the set of local minima of $H_\tau(x_\tau, IP_\tau)$ with respect to IP_τ that yields the optimal second order quantity as

$$Q_\tau^* = \begin{cases} -x_\tau & \text{If } x_\tau \le IP_\tau^0\\ 0 & \text{If } IP_\tau^{i-1} < x_\tau \le a^i \quad i = 1, \ldots, n\\ IP_\tau^i - x_\tau & \text{If } a^i < x_\tau \le IP_\tau^i \quad i = 1, \ldots, n\\ 0 & \text{If } IP_\tau^n < x_\tau. \end{cases} \tag{5}$$

It follows from this policy that $\rho_\tau(x_\tau)$ in (P1.1) is piecewise convex. To solve (P1.2), rather than analyze $\rho_\tau(x_\tau)$ directly, we approach the problem by first considering the two functions introduced below:

$$j_\tau(x_\tau) = \begin{cases} H_\tau(x_\tau, 0) & \text{If } x_\tau \le IP_\tau^0\\ H_\tau(x_\tau, IP_\tau^i) & \text{If } IP_\tau^{i-1} < x_\tau \le IP_\tau^i \quad i = 1, \ldots, n\\ H_\tau(x_\tau, x_\tau) & \text{If } IP_\tau^n < x_\tau, \end{cases} \tag{6}$$

$$k_\tau(x_\tau) = \begin{cases} H_\tau(x_\tau, 0) & \text{If } x_\tau \le IP_\tau^0\\ H_\tau(x_\tau, x_\tau) & \text{If } x_\tau > IP_\tau^0. \end{cases}$$

It can be seen that $\rho_\tau(x_\tau) = \min\{j_\tau(x_\tau), k_\tau(x_\tau)\}$ and that it will yield the optimal policy defined in (5). Except in the boundary cases where $j_\tau(x_\tau)$ and $k_\tau(x_\tau)$ are set to be identical, $k_\tau(x_\tau)$ represents the expected cost if the second order is not placed, and, $j_\tau(x_\tau)$ represents the minimum expected cost given that an order must be placed to bring the inventory position to the smallest IP_τ^i not less than x_τ.

We then proceed to define

$$\begin{aligned} j_1(x_1) &= \sum_{t=1}^{\tau-1} b_t \int_{x_1}^{\infty} (\xi_{1,t} - x_1) f_{1,t}(\xi_{1,t})\, d\xi_{1,t}\\ &\quad + \sum_{t=1}^{\tau-1} h_t \int_0^{x_1} (x_1 - \xi_{1,t}) f_{1,t}(\xi_{1,t})\, d\xi_{1,t}\\ &\quad + \int_0^{\infty} j_\tau(x_1 - \xi_{1,\tau-1}) f_{1,\tau-1}(\xi_{1,\tau-1})\, d\xi_{1,\tau-1},\\ k_1(x_1) &= \sum_{t=1}^{\tau-1} b_t \int_{x_1}^{\infty} (\xi_{1,t} - x_1) f_{1,t}(\xi_{1,t})\, d\xi_{1,t}\\ &\quad + \sum_{t=1}^{\tau-1} h_t \int_0^{x_1} (x_1 - \xi_{1,t}) f_{1,t}(\xi_{1,t})\, d\xi_{1,t}\\ &\quad + \int_0^{\infty} k_\tau(x_1 - \xi_{1,\tau-1}) f_{1,\tau-1}(\xi_{1,\tau-1})\, d\xi_{1,\tau-1}. \end{aligned}$$

Here, the first two summation terms for each of $j_1(x_1)$ and $k_1(x_1)$ represent, respectively, the expected backorder and holding costs in periods 1 through $\tau - 1$. Thus, it follows that

$$TC_I(Q_0) = \min\{c_0 Q_0 + j_1(Q_0), c_0 Q_0 + k_1(Q_0)\}, \qquad (7)$$

whose derivation has taken advantage of (3) and $\rho_\tau(x_\tau) = \min\{j_\tau(x_\tau), k_\tau(x_\tau)\}$.

In general a line search would be required to obtain all admissible breakpoints IP_τ^i $(i = 0, \ldots, n)$ before the choice in (7) can be made. Alternatively, we may put restrictions on the specification of the problem to guarantee that $H_\tau(x_\tau, IP_\tau)$ is sufficiently well behaved in IP_τ to yield one local minimum. This results in the following theorem.

Theorem 1.

(1) *If $L = 0$, or $L > 0$, $b_\tau = \cdots = b_{\tau+L-1} = 0$, then $H_\tau(x_\tau, IP_\tau)$ is convex in IP_τ and $TC_I(Q_0)$ is convex in Q_0.*

(2) *If $L = T - \tau + 1$, $b_\tau = \cdots = b_{\tau+L-2} = 0$, and $b_{\tau+L-1} > 0$, then $H_\tau(x_\tau, IP_\tau)$ is convex in IP_τ and $TC_I(Q_0)$ is convex in Q_0.*

(3) *If $L > 0$, $b_\tau = \cdots = b_{\tau+L-2} = 0$, $b_{\tau+L-1} > 0$, $h_{\tau+L} = \cdots = h_{T-1} = 0$, and the monotone likelihood ratio property (MLRP) holds for $f_{\tau,T}(IP_\tau)/f_{\tau,\tau+L-1}(IP_\tau)$ over $[0, \infty)$, then $H_\tau(x_\tau, IP_\tau)$ has at most one local maximum and one local minimum with respect to IP_τ.*

(4) *If $c_\tau + \sum_{t=\tau}^{\tau+L-1} b_t - c_u \leq 0$ and $\xi_\tau, \ldots, \xi_{\tau+L-1}$ have independent PF_2 densities, then $H_\tau(x_\tau, IP_\tau)$ is unimodal in IP_τ and $TC_I(Q_0)$ is convex in Q_0.*

(5) *If $\xi_\tau, \ldots, \xi_{\tau+L-1}$ have independent PF_3 densities, then $H_\tau(x_\tau, IP_\tau)$ has at most one local minimum and one local maximum with respect to IP_τ.*

Theorem 1 identifies five specifications, each of which guarantees that (4) has at most two roots, one of which is the minimum. Interestingly, the antecedents of Theorem 1 Part (2) can be traced to the work of Barankin (1961). In his model, an "emergency" order is placed and delivered at the beginning of the season so that $\tau = 1$, whereas the second order arrives at the end of the season, so that L is effectively equal to T. When this is the case it is easy to verify that under the conditions of Part (2) the problem remains separable, and that it is convex in each of the two variables. This is also the case in the more recent two-order-opportunity model of demand management of Iyer et al. (2003). In this model, an initial order, corresponding to Q_0 is placed for delivery at the start of the season. Subsequently, there is an information phase in which the demand forecast for the season is revised. This revised information is considered in placing the second order, Q_1 at $\tau = 1$. And, this information is concurrently used to split demand into two groups: one that will be satisfied from Q_0, and the other that will be satisfied from Q_1, and possibly, the unused portion of Q_0. In this sense, this model also has an effective leadtime of $L = T$. As will be discussed later, incorporating forecast revision, based on information from the preseason, can be easily accommodated in our construct of the problem.

Exploiting the benefit of updated information from the preseason also underlies the work of Donohue (2000), whose model is a one-product variant of the model of Fisher and Raman (1996); a similar model is also considered by Jones et al. (2001). In these models $L = 0$ and $\tau = 1$, so that demand is either satisfied from inventory on hand or it is lost. Hence, as articulated in Theorem 1 Part (1), this yields convexity in $IP_1 = Q_0 + Q_1$.

The last three parts of Theorem 1 accommodate the case when L is between 0 and T, but place increasingly more restrictive regularity conditions as the modeling environment is enriched. Part (3) addresses those scenarios in which backorder costs are not time dependent; the densities must exhibit the frequently used monotone likelihood ratio property (Karlin and Rubin 1956). When backorder costs are time dependent, Part (4) accommodates the case when it is always more economical to use a unit from the second order to meet a backorder than to let it be lost, but the densities must exhibit the specific variation diminishing property of the PF_2 class. Analogously, under arbitrary economic parameters, the densities must exhibit the more restrictive variation diminishing property of the PF_3 class, which is a subset of the PF_2 class. The development of PF_n densities is due to Schoenberg (1951), and its initial application to stochastic inventory models appears to be due to Karlin (1958, 1968). Recently, Porteus (2002) provided an intuitive explanation of this family of densities. Importantly, PF_2 densities have the monotone likelihood ratio property, and PF_2, PF_3, and MLRP densities include the classes of truncated normal and gamma families.

It is important to reiterate that our model is a generalization of the model considered by Fisher et al. (2001) that allows c_0 to be different than c_τ and allows the backorder cost during the leadtime to be time-dependent. Fisher et al. (2001) focused on developing algorithmic solutions to the problem because they observed that this problem is not convex in the two decision variables. In contrast, we have shown that under regularity conditions that accommodate the case of normal demand considered by them, the problem can be reduced to a convex optimization problem in one variable.

Theorem 1 establishes the conditions under which the problem is well behaved for a given τ. We now complement these structural results by examining the impact of varying the timing of the second order. To this end, consider the case in which the retailer may be able to negotiate different trade terms for potential values for τ. Then, as in Milner and Kouvelis (2002) and Fisher et al. (2001), for every viable value of τ, we would solve problem (P1.2) and then choose the τ that gives the lowest expected cost. Because this more complex structure can be shown by numerical examples not to be well behaved, a line search must be used to find the optimal τ, as in Example 1 in the online appendix. This representative example also tracks how changes in τ affect the optimal choice of the two order quantities or more.

Although we have formulated our model without allowing for forecast revision, it can easily accommodate forecast revision in the preseason as well as additional updates based on observed demands under appropriate regularity conditions. Let S be the vector of information from the preseason and from demand during periods 1 through $\tau - 1$ that is equivalent to a sufficient statistic available at the beginning of period τ; this statistic is used to determine posterior distributions of demand. The key to the analysis is to recognize that if $\rho_\tau(x_\tau)$ conditional on the sufficient statistic is convex, then, under the expectation over the distribution of S, it will remain convex and, therefore, $\rho_1(x_1)$ is convex by (3) (see Scarf 1959). Under the conditions of Theorem 1, Parts (1) and (2), the convexity of $\rho_\tau(x_\tau)$ is always assured so any updating method can be used. Whereas for Part (4) we must assure that the revised distributions remain in the PF_2 class. This is guaranteed if forecast revision is executed using conjugate priors (see, for example, Berger 1985), as illustrated by Example 2 in the online appendix.

To summarize, so far we have shown that solving (P1.1) reduces to sequentially determining the initial order quantity after determining the second order quantity for all possible states of x_τ at the beginning of period τ. If finding the target level for the second order entails minimizing a unimodal function, determining the initial order quantity reduces to finding the unique minimum of a convex function. Under the same regularity conditions, our model can be generalized to incorporate forecast revisions. Moreover, in the discussion in this section we have shown how our model generalizes many variants of the recent and early literature on newsvendor models with two purchase opportunities. We have captured those models in which either the second order is placed at the start of the season or at a predetermined point during the season. Hence, adapting the terminology of Milner and Kouvelis (2002), the timing of the second order is static whereas the choice of the order quantity is dynamic. We next consider a variant of the model in which both the second order quantity and its timing are determined dynamically.

3. Extended Model

Using results from the previous section as building blocks, we consider our most general setting depicted in Figure 1. Not only does this setting preserve the flexibility of determining the second order quantity dynamically, it offers the additional flexibility of determining dynamically when the second order is placed. Hence, if demand is higher than anticipated when the first order was placed, it is likely that the second order will be placed earlier in the season for a larger quantity. In contrast, if demand is lower than anticipated, the second order will be placed later in the season for a smaller quantity. It is the goal of this section to formulate and analyze the structure of this problem.

To proceed with the generalization, we amend the dynamics of the basic model. As in the basic model, in the extended model we start the season with Q_0 units in stock so that $x_1 = Q_0$. Then at the beginning of each period t, $t = 1, \ldots, T - L + 1$, we determine whether to place the second order at unit cost c_t. If the second order is not placed, we accept all demand in period t; that in excess of x_t is backordered for guaranteed delivery at some time in the future. However, if the second order is placed in period t, we continue to accept demand until the inventory position reaches 0, with the understanding that all backorders will be cleared when the second order is received. All subsequent demand is considered lost at a unit penalty cost c_u.

Because the dynamics of the extended model are similar to those of the basic model, adapting its notation facilitates the analysis. To proceed we need to amend and append the following notation.

c_t: The unit purchase cost for the second order if placed at the beginning of period t; $t = 1, \ldots, T-L+1$.

$g_t(x_t, y_t)$: The expected cost in periods t through T with a beginning inventory level x_t and a beginning inventory position y_t after the second order is placed at the beginning of period t, excluding the purchase cost of the two orders (analogous to $g_\tau(x_\tau, y_\tau)$ in the basic model); $t = 1, \ldots, T-L+1$.

$\rho_t(x_t)$: The minimum expected cost in periods t through T with a beginning inventory level x_t when the second order opportunity is still available, excluding the purchase cost for the first order; $t = 1, \ldots, T-L+1$.

$k_t(x_t)$: The minimum expected cost in periods t through T when the second order is not placed at the beginning of period t, including the expected purchase cost for the second order in a future period; $t = 1, \ldots, T-L+1$. For $t = 1, \ldots, T-L$, we have

$$k_t(x_t) = b_t\int_{x_t}^{\infty}(\xi_t - x_t) f_t(\xi_t)\,d\xi_t + h_t\int_0^{x_t}(x_t-\xi_t)f_t(\xi_t)\,d\xi_t + \int_0^{\infty}\rho_{t+1}(x_t-\xi_t)f_t(\xi_t)\,d\xi_t.$$

$j_t(x_t)$: The minimum expected cost in periods t through T when the second order is placed at the beginning of period t, including the purchase cost for the second order; $t = 1, \ldots, T-L+1$ ($k_{T-L+1}(x_{T-L+1})$ and $j_{T-L+1}(x_{T-L+1})$ are analogous to k_τ and j_τ in the basic model, respectively).

$$j_t(x_t) = \min_{IP_t \geq x_t}\{H_t(x_t, IP_t) = c_t(IP_t - x_t) + g_t(x_t, IP_t)\}. \quad (8)$$

Based on the notation above, we can formulate the decision problem for the second order opportunity at the beginning of period t. Because there are two choices at the beginning of period t: either no order is placed incurring the cost $k_t(x_t)$ or an order is placed incurring the cost $j_t(x_t)$, we have

$$\rho_t(x_t) = \min\{k_t(x_t), j_t(x_t)\}. \quad \text{(P2.1)}$$

With a beginning inventory level x_t, an order is placed in period t only when $k_t(x_t) \geq j_t(x_t)$. Similarly, the decision problem for the first order quantity prior to the start of the season is formulated as

$$\min_{Q_0 \geq 0}\{TC_{II}(Q_0) = c_0Q_0 + \rho_1(x_1 = Q_0)\}. \quad \text{(P2.2)}$$

To proceed with the analysis, notice that because period $T-L+1$ is the last period when an order can be placed, it follows that the optimal order quantity determined essentially from (8) for the second order placed at the beginning of period t is given by the equivalent of (5). Hence, it follows that the optimal policy for the second order quantity is generally quite complex, so that appealing policies of the form (s, S) do not apply optimally; that is, ordering up to S if the inventory level is below s cannot be optimal in general. This is despite the fact that under the conditions of Theorem 1, Parts (2) and (4), $\rho_t(x_t)$ is convex in x_t, which is sufficient for the optimality of such policies in the basic model. Unfortunately, although these conditions are sufficient to guarantee that, for a small enough x_t, it is optimal to place an order in period t, they are not strong enough to sufficiently smooth the cost function $k_{t+1}(x_{t+1})$ to assure that $\rho_t(x_t)$ is well behaved in x_t in general. In particular, we are able to generate robust counter examples with $L > 0$ that show that $k_t(IP_t^*) > j_t(IP_t^*)$ (see the online appendix). As a consequence, in any meaningful policy, $j_t(x_t)$ and $k_t(x_t)$ may cross an even number of times, making (s, S) policies sub-optimal.

Although we are in general unable to establish that the optimal policy structure is well defined, under appropriate regularity conditions the optimal ordering policy is of the (s, S) class as stated in the following theorem, which may be viewed as an amalgamation of the first three parts of Theorem 1.

Theorem 2. *If $c_t + h_t \leq c_{t+1}$ for all t and*

(1) If $L = 0$ and demand $\xi_1, \ldots, \xi_T$ have PF_2 densities, then the (s_t, S_t) policy is the optimal decision rule for whether or not to place the second order when the second order opportunity is available.

(2) *If* $L > 0, b_1 = \cdots = b_{T-1} = 0, b_T \geq 0$, *and demand* $\xi_1, \ldots, \xi_T$ *have* PF_2 *densities, then the* (s_t, S_t) *policy is the optimal decision rule for whether or not to place the second order when the second order opportunity is available.*

Under conditions of Theorem 2, we can show that, in period t, $j_t(x_t) \geq k_t(x_t)$ for $x_t \geq IP_t^*$, $j_t(x_t) < k_t(x_t)$ for small enough x_t, and that $j_t(x_t)$ and $k_t(x_t)$ cross only once. The intersection point of $j_t(x_t)$ and $k_t(x_t)$ is the order-triggering point s_t, and the minimizer, IP_t^*, from (8) is the order-up-to level S_t. These conditions also imply that $TC_{II}(Q_0)$ in (P2.2) is well behaved in Q_0. This leads to the following theorem.

Theorem 3. *Under conditions of Theorem* 2 *the optimal* Q_0 *in* (P3.2) *is unique.*

Although variants of the basic model have appeared in the literature, we are not aware of any work that has established the optimality of (s, S) policies for the case when the timing of the second order is dynamically determined. Interestingly, Milner and Kouvelis (2002, 2005) advocated the use of this policy in the lost-sales version of this problem. They presented results from computational experience in finding heuristically determined values of s and S for the case when the demand (rate) follows a Brownian motion. In contrast, under the conditions of Theorem 2, determining the optimal choices for Q_0 and the second order reduces to a series of one-variable optimization problems, each with a unique solution.

On a technical level, as illustrated in the preceding section, the basic model can accommodate forecast revision in the preseason as well as additional updates based on the observed demands under our regularity conditions because the cost function $\rho_\tau(x_\tau)$, conditional on the sufficient statistic, is convex. We suspect that the extended model, even under conditions of Theorem 2, cannot accommodate forecast revision because the cost function $\rho_t(x_t)$ is only unimodal, and not necessarily convex. On a managerial level, the extended model provides the additional operational flexibility of determining the timing of the second order dynamically, whereas in the basic model the timing of the second order is chosen at the beginning of period 1. We present Example 3 in the online appendix to show that the value of the operational benefits brought by the flexible placement of the second order appears limited.

4. Conclusions

In this paper we have conducted an extensive analysis of the inventory management problem faced by a retailer of a short-lived product. In its traditional representation, such a retailer may be viewed as a newsvendor who makes a stocking decision before the start of the selling season. Our problem, depicted in Figure 1, differs from that of the newsvendor because here the decision maker has an opportunity to place a second order to fine tune the preseason stocking decision made in the traditional newsvendor setting. As a consequence more refined decisions may be made that allow opportunities to incorporate forecast revision and/or the dynamic evolution of inventory that is modulated by the evolution of demand.

Although it is the case that the antecedents of our model may be traced to early work by Barankin (1961), the modeling framework is more general. It includes as a special case the single-product variant of the model of Fisher and Raman (1996) that is considered in detail by Donohue (2000) and Jones et al. (2001). Our framework generalized these models by allowing the second order to be delivered in midseason.

In this class of models, because all stocking decisions are made before the start of the season, its primary benefit is that demand forecasts may be revised when there is a belief that the preseason provides opportunities to obtain more accurate forecasts. There is another class of models in which the second order is placed at a predetermined time during the season for delivery at a later point of the selling season. Thus, in addition to using observed demand to further refine the demand forecast, there is an opportunity to use the second order to move the target inventory up or down to reflect higher or lower observed demand rates. Notably, this class includes the model of Fisher et al. (2001). For this setting, our framework provides a complete characterization of the optimal policy structure under significantly more general conditions than those considered by them.

In these classes of problems the timing of the second order is given, so the focus is on determining the optimal initial and second order quantities. In the final class of models that we consider, the timing of the second order is determined dynamically after taking into account the evolution of demand. In contrast to the previous two classes, the optimal policy structure is generally complex and so not tractable analytically. For example, we demonstrate that theoretically attractive (s, S) policies are not necessarily optimal for this class of problems.

While elucidating the structure of the optimal policy for problems that fit our framework is an important technical contribution of our work, these technical results are complemented by computational work. A technical implication of our numerical examples, reported in the online appendix, is that they can be used to demonstrate that the problem structure is not well behaved for the basic model in which the predetermined timing or leadtime of the second order is optimized.

In addition to such technical insights, our computational work has qualitative implications. A series of representative solved experiments suggest that much of the benefit of the second order can be realized even when the timing of the second order is predetermined prior to the start of the season. In other words, endogenizing the timing of the second order has relatively small savings. Interestingly, our numerical examples suggest that, in the expected sense, the aggregate order quantity is relatively insensitive to the timing of the order. This allows us to conclude that the service delivered to consumers is relatively insensitive to the optimal timing of the two order quantities. However, the same examples demonstrate that the relative costs incurred are sensitive to carefully choosing the time when the second order is placed.

Although we have focused on the impact of a second order on managing inventories for products with short selling seasons, we conclude by identifying other directions for managing inventories in these settings. Notice that in our models the second order opportunity was only used to increase the inventory level. However, it can also be beneficial to decrease the current inventory as in Cachon and Kok (2007), Kouvelis and Gutierrez (1997), and Petruzzi and Dada (2001). Consistent with the literature and the practice of quick response, the addition of the second order opportunity has benefits to the retailers managing inventory levels for short selling season products, which is our perspective of this research. Alternatively, one may show that it can bring benefits to the upstream suppliers through appropriately designed mechanisms that allow for Pareto improvements of all involved parties, as demonstrated by Cachon (2004), Ferguson et al. (2005), Iyer and Bergen (1997), and Taylor (2006).

Electronic Companion

An electronic companion to this paper is available on the *Manufacturing & Service Operations Management* website (http://msom.pubs.informs.org/ecompanion.html).

Acknowledgments

The authors thank Gérard Cachon, Garrett van Ryzin, the associate editor, and two anonymous referees for their invaluable insights and suggestions.

References

Barankin, E. W. 1961. A delivery-lag inventory model with an emergency provision (the single period case). *Naval Res. Logist. Quart.* **8** 285–311.

Berger, J. O. 1985. *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer-Verlag, New York.

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

Cachon, G., G. Kok. 2007. Implementation of the newsvendor model with clearance pricing: How to (and how not to) estimate a salvage value. *Manufacturing Service Oper. Management* **9**(3) 276–290.

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

Eppen, G. D., A. V. Iyer. 1997. Backup agreements in fashion buying—The value of upstream flexibility. *Management Sci.* **43**(11) 1469–1484.

Ferguson, M., G. DeCroix, P. Zipkin. 2005. Commitment decisions with partial information updating. *Naval Res. Logist.* **52**(8) 780–795.

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

Fisher, M., K. Rajaram, A. Raman. 2001. Optimizing inventory replenishment of retail fashion products. *Manufacturing Service Oper. Management* **3**(3) 230–241.

Iyer, A., M. Bergen. 1997. Quick response in manufacturer-retailer channels. *Management Sci.* **43**(4) 559–570.

Iyer, A. V., V. Deshpande, Z. Wu. 2003. A postponement model for demand management. *Management Sci.* **49**(8) 983–1002.

Jones, P. C., T. Lowe, R. D. Traub, G. Kegler. 2001. Matching supply and demand: The value of a second order chance in producing hybrid seed corn. *Manufacturing Service Oper. Management* **3**(2) 122–137.

Karlin, S. 1958. One stage model with uncertainty. *Studies in the Mathematical Theory of Inventory and Production*. Stanford University Press, Stanford, CA.

Karlin, S. 1968. *Total Positivity*, Vol. 1. Stanford University Press, Stanford, CA.

Karlin, S., H. Rubin. 1956. The theory of decision procedures for distributions with monotone likelihood ratio. *Ann. Math. Statist.* **27**(2) 272–299.

Kouvelis, P., G. J. Gutierrez. 1997. The newsvendor problem in a global market: Optimal centralized and decentralized control policies for a two market stochastic inventory system. *Management Sci.* **43**(5) 571–585.

Milner, J. M., P. Kouvelis. 2002. On the complementary value of accurate demand information and production and supplier flexibility. *Manufacturing Service Oper. Management* **4**(2) 99–113.

Milner, J. M., P. Kouvelis. 2005. Order quantity and timing flexibility in supply chains: The role of demand characteristics. *Management Sci.* **51**(6) 970–985.

Petruzzi, N., M. Dada. 2001. Information and inventory recourse for a two-market, price-setting retailer. *Manufacturing Service Oper. Management* **3**(3) 242–263.

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

Schoenberg, I. S. 1951. On Pólya frequency function. *J. d'Analyse Mathématique* **I** 331–374.

Scarf, H. 1959. Bayes solutions of the statistical inventory problem. *Ann. Math. Statist.* **30**(2) 490–508.

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