A DETERMINISTIC SINGLE-ITEM INVENTORY MODEL WITH SELLER HOLDING COST AND BUYER HOLDING AND SHORTAGE COSTS

AMIHAI GLAZER

University of California, Irvine, California

REFAEL HASSIN

Tel Aviv University, Tel Aviv, Israel

(Received October 1982; revisions received November 1983, September 1984, May 1985; accepted December 1985)

This paper considers the optimal deterministic inventory policy for a monopolist. The novel element concerns the consumers. Each has a time at which he would most prefer to obtain the good; he incurs a shortage cost if he buys it after that time, and a holding cost if he obtains it before. The firm must consider that any change in its policy will affect the quantity of goods it can sell and the maximum price it can charge. We find that the seller's optimal policy can take only one of three possible forms: sell the goods throughout each cycle; sell the goods continuously during an interval that is shorter than an inventory cycle; or sell the goods only at the instant each cycle begins.

The classical deterministic inventory model (see Hadley and Whitin 1963) adopts two approaches to shortages. The first is lost sales, wherein customers who wish to purchase the good at a time when the firm is out of stock never make the purchases. The second approach allows for backorders. The firm postpones sales to customers who wish it at a time of shortages, but in doing so the firm incurs a penalty for each unit backordered, and a variable cost that increases with the length of time customers must wait.

Under the assumption of lost sales, profit maximization requires that either no sales be made, or no stockouts occur. Under the assumption of penalties for backorders, profit maximization requires stockouts if the variable shortage cost is finite and positive. If this cost is zero, then either there will be no stockouts, or the backorders will not be filled.

But buyers, not only sellers, can hold inventory. A buyer who prefers to purchase a good at some particular time might be willing to buy it before that time if the price and his inventory carrying costs are sufficiently low. Similarly, he might still be willing to buy it some time later than his most preferred time if his shortage costs are small. A manufacturer, for example, committed to producing some output by June, might be willing to pay \$300 for steel delivered to him in May, but only \$200 for steel delivered in February and only \$100 for steel delivered in July. The timing of purchases by customers is thus subject to some control by the seller: he can lower his total inventory carrying costs by inducing some customers to purchase the good early in each of the seller's inventory cycles.

Though two recent papers consider models in which consumers can postpone purchases (Bucovetsky 1983, Sobel 1984), no paper that we know of considers consumers who may purchase the good either before or after that instant at which they value it most highly.

Our model does not assume any explicit stockout cost. Instead, we begin by assuming that each buyer has a preferred time for receiving the good, and that he incurs a shortage or a holding cost if the good is not then delivered. The profit-maximizing firm, however, must take these costs into account when determining its optimal policy. We shall demonstrate that, in general, the optimal policy requires a stockout period within each replenishment cycle. Nevertheless, all customers will be served, and under some circumstances the firm will backorder all of the demand while choosing a replenishment cycle of finite length (this type of solution is frequently encountered in the real world, but is not explained by the classical models).

The main contribution of this paper is to present a fresh approach to the analysis of inventory systems, an approach that can be used in models other than that given in this analysis. Its main substantive contribution is to show that the seller's optimal policy can take only three possible forms: 1) sell the goods throughout each cycle, 2) sell the goods continuously

Subject classification: 334 buyer holding and shortage costs.

0030-364X/86/3404-0613 \$01.25

during an interval that is shorter than an inventory cycle, and 3) sell the goods only at the instant each cycle begins.

In the following sections we demonstrate this assertion and give the conditions under which each of the three policies is optimal. For expository purposes we distinguish between two circumstances: the buyer's shortage costs are less than his holding costs or greater than them. These two cases are treated in Sections 2 and 3, respectively. In fact, the optimal policies described in Section 3 are a subset of those discussed in the former section. Section 4 ties together the preceding sections.

1. Assumptions

All inventory models must assume that the seller has some monopoly power. The alternative assumption—perfect competition—is inconsistent with inventory holding. By definition, each firm in a perfectly competitive industry can sell whatever quantity it wishes at the going market price; a firm could sell *all* of its goods immediately upon receipt of a new shipment and would have no need to carry any inventory. We assume accordingly that the firm is a monopolist that desires to maximize profits. Let h be its cost of holding one unit of inventory for one unit of time. Its marginal cost of obtaining a unit of the good is c. The fixed cost of placing an order of any size is A. The firm chooses a fixed price, p, and an inventory cycle of length T. For convenience, we shall usually speak of one such cycle that begins at time 0 and ends at time T.

A customer's valuation of the good is a function of the time at which he purchases it. Let t_p be the time at which some customer would most prefer to obtain the good; he would be willing to pay a price w for a good obtained then. In purchasing a good before that time, the buyer incurs a holding cost of i per unit of time, so that his reservation price, $v(t)$, at time t is $w - i(t_p - t)$ for $t < t_p$. In purchasing a good after time t_p, the buyer incurs a shortage cost of s per unit of time, so that $v(t)$ is $w - s(t - t_p)$ for $t > t_p$. All buyers have the same values of w, s and i; they differ only in their preferred purchase times. These reservation prices are illustrated in Figure 1.

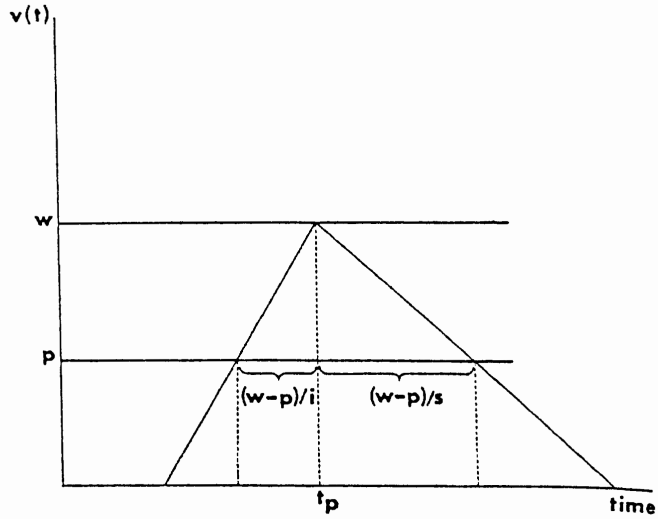

Figure 1. Consumer reservation prices.

A buyer's w-time is defined as the value of t that satisfies $v(t) = w$; the buyer depicted in Figure 1, for example, has a w-time of t_p. Each buyer is characterized by a different w-time, which is exogenously fixed; his choices are whether and when to purchase the good.

As in the Economic Order Quantity (EOQ) model, we assume that the rate of customer arrivals is constant; that is, the number of buyers with a w-time in the interval $(t, t + dt)$ is λdt for all values of t and of dt.

The firm treats the values of λ, s, h, and i as fixed. It must decide on the price p (with $p \leq w$); on the length, T, of an inventory cycle; and on the times at which to sell the good.

We proceed with the analysis as follows. Proposition 1 states that under the firm's optimal policy all customers will purchase the good. The next three propositions concern the timing of sales. If $s < i$, sales will be made either continuously during some interval (the interval can be shorter than the length of a cycle), or else only at times $0, T, 2T\ldots$ (Proposition 2). If $s > i$, sales will be made either continuously throughout each cycle, or only at discrete instants within it (Proposition 3). Proposition 4 further specifies the form of the optimal policy when $s > i$: discrete sales can occur only at times $0, T, 2T\ldots$.

Proposition 1. *Suppose the firm finds it profitable to make some sales. Then it will choose a price and sales intervals so that all customers purchase the goods.*

Proof. Let S denote the set of points at which the firm sells the goods. Clearly, if S is nonempty, $0, T, 2T\ldots$ are in S; for otherwise, the firm could save inventory costs by postponing an order time to the next sales point. Let p be the profit-maximizing price. Clearly, $p \leq w$, for otherwise no sales are made.

Suppose, contrary to our assertion, that for some nonempty interval (t_1, t_2), all customers with w-times in that interval find that $v(t) < p$ for all $t \in S$. Let this interval (t_1, t_2) be a maximal interval: that is, it is not properly contained in another interval with the same property. Clearly (t_1, t_2) does not intersect with S, for otherwise customers with w-times in this interval would, contrary to the assumption, purchase the good at their w-times. Define t_2' as the smallest value of t greater than or equal to t_2 that lies in S. By construction, the intersection of interval (t_2, t_2') with S is the null set.

Suppose first that $t_2' < T$, and consider a modified policy in which S is extended to include the interval $[t_1 + t_2' - t_2, t_2']$. Note that customers with w-times before t_2 refused to make a purchase at t_2'. Therefore, because of their shortage costs, customers with a w-time before t_1 will not postpone their purchases to somewhere in $[t_1 + t_2' - t_2, t_2']$. The change in the sales policy therefore affects only two classes of customers. Customers with w-times in $[t_1 + t_2' - t_2, t_2']$ will find it worthwhile to purchase the good in $[t_1 + t_2' - t_2, t_2']$. And, if $p < w$, some customers who previously purchased the good at t_2' will prefer to buy it earlier, within $[t_1 + t_2' - t_2, t_2']$. Both effects increase profits: the first since for every $t \leq t_2'$ the value of $c + ht \leq c + ht_2' \leq p$; the second by reducing the firm's inventory costs. Thus when $t_2' < T$, the initial policy is suboptimal.

Suppose next that $t_2' = T$, so that no sales are made in the interval (t_1, T). Consider a modified policy that is identical to the initial one, except that the length of a cycle is shortened from T to $T - \Delta$, where $\Delta < t_2 - t_1$. The behavior of customers with w-times in the interval $(0, t_1)$ is the same under the two policies, and therefore the inventory carrying costs and revenue earned from these customers are identical. We shall consider the behavior of customers with w-times in (t_1, T), and in $(t_1, T - \Delta)$ under the two policies.

Initially, no sales are made to customers with w-times in the interval (t_1, t_2), and customers with w-times in (t_2, T) make purchases at time T, when a new shipment arrives. Under the modified policy, with a new shipment at $T - \Delta$, customers with w-times in $(t_1, t_2 - \Delta)$ make no purchases, and those with w-times in $(t_2 - \Delta, T - \Delta)$ make the purchases at time $T - \Delta$. Therefore, under the two policies the same pattern of sales holds during the interval $(0, t_1)$, under neither policy are sales made from time t_1 until the end of a cycle, and the number of goods purchased when a new shipment arrives is the same under the two policies. The modified policy therefore obtains the same profit per cycle as the initial policy, but in a shorter period, so that it yields a greater profit per unit time.

We conclude that a policy that makes sales to some, but not all, customers is not optimal.

Since sales are made to all customers, the firm's cost of purchasing the goods per unit time, λc, is a fixed cost and can be ignored in subsequent analyses. We therefore let $c = 0$.

Corollary. *The length of a no-sales interval is bounded by* $(w - p)(1/i + 1/s)$.

Proof. From Proposition 1, all customers purchase the good. A person with a w-time of t_p will purchase the good if $v(t) \geq p$, that is, in the interval $(t_p - (w - p)/i, t_p + (w - p)/s)$. Thus, if all customers make purchases, the length of a no-sales interval can be no greater than $(w - p)(1/i + 1/s)$.

The next two propositions concern the timing of sales.

Proposition 2. *If $s < i$ and the firm finds it profitable to make sales at times t_1 and t_2 $(0 \leq t_1 < t_2 < T)$, then it will find it even more profitable to make sales all during the interval (t_1, t_2).*

Proof. Consider a no-sales period (t_1, t_2) with $0 \leq t_1 < t_2 < T$, for which sales are made at t_1 and t_2. Of the customers with w-times in (t_1, t_2), a fraction $s/(i + s)$ purchase the good at t_1, and a fraction $i/(i + s)$ purchase it at t_2. If, instead, sales are continuously made during the interval, these customers will buy the good at their w-times, which are evenly distributed over the interval. When $s < i$, this alternative entails a decrease in holding costs, even if p is maintained at its previous level.

Note that if $t_2 = T$, optimality need *not* require continuous sales during the interval (t_1, t_2): the previous proof would not be valid, since delaying the sales to time T would decrease the firm's inventory costs. Therefore, even when $s < i$, it is possible that a no-sales period (t_2, T) will be part of the optimal solution.

Proposition 3. *If $s > i$, the firm makes sales either throughout each cycle, or else at discrete points within each cycle.*

Proof. If $p = w$, each customer will purchase the good at his w-time. By Corollary 1, the no-sales interval is empty and sales will be made throughout each cycle. Consider next the case in which $p < w$, so that buyers might be willing to purchase the good either before or after their w-times. Suppose, contrary to the assertion in Proposition 3, that there exists some interval within each cycle during which sales are made throughout. Consider such a closed interval $[t_1, t_2]$ with $0 \leq t_1 < t_2 \leq T$ and $t_2 - t_1 \leq (w - p)(1/i + 1/s)$. The only persons who would make purchases inside the open interval (t_1, t_2) are those with w-times in it. By refusing, however, to make sales inside this interval, the firm would induce a fraction $s/(i + s)$ of these persons to purchase the good at time t_1, and a fraction $i/(i + s)$ to purchase it at time t_2. If $s > i$, such a change would reduce the firm's inventory holding costs even if p is maintained at its previous level.

2. Optimal Policy when $s < i$

This section analyzes the optimal inventory policy when $s < i$. (The other case is discussed in Section 3.) Since all cycles are identical, we need concern ourselves with only one of them, which we shall say commences at time 0 with a new inventory shipment, and ends at time T. Proposition 2 and its discussion showed that if $s < i$, the firm's optimal policy is to make sales throughout an interval $[0, T_D]$, $0 \leqslant T_D \leqslant T$, and to make no sales in the interval (T_D, T).

Corollary 1 states that for a given price, p, the length of the no-sales interval can be no greater than $(w - p)(1/i + 1/s)$, so that the value of T_D can be no less than $T - (w - p)[(1/i) + (1/s)]$. If T_D is greater than that value, some customers will purchase the good inside the interval $(T - (w - p)[(1/i) + (1/s)], T)$ rather than at its end points. Recall that T represents the beginning of a new cycle and the receipt of a new shipment, so that the former sales pattern entails higher inventory carrying costs than does the latter.

Profit maximization therefore requires that

$$T_D = T - (w - p)\left[\frac{1}{i} + \frac{1}{s}\right]. \tag{1}$$

The firm's revenue per cycle is λpT. Its inventory holding expenses are a function of when customers make their purchases. Buyers with w-times in the interval $(0, T_D)$ purchase the good at their w-times; the number of such buyers is λT_D. For buyers with w-times in the interval $(T_D, T_D + (w - p)/i)$, the value of $v(T_D)$ is greater than of $v(T)$, so that they purchase the good at time T_D; the number of such buyers is $\lambda(w - p)/i$. Buyers with w-times in the interval $(T_D + (w - p)/i, T)$, which is equivalent to the interval $(T - (w - p)/s, T)$, purchase the good at time T; there are $\lambda(w - p)/s$ such buyers.

The firm's average rate of profit is therefore

$$\pi = -\frac{A}{T} + \frac{\lambda}{T}\left[pT - \frac{hT_D^2}{2} - \frac{hT_D(w - p)}{i}\right]. \tag{2}$$

Substitute Equation 1 in Equation 2 and differentiate with respect to T and p to find the necessary first order conditions for the maximization of profits:

$$p = w - \beta T \tag{3}$$

and

$$T = \sqrt{\gamma} \cdot \sqrt{2A/\lambda h}, \tag{4}$$

where

$$\beta \equiv \frac{(1/s) - (1/h)}{(1/s)^2 - (1/i)^2}$$

and

$$\gamma \equiv \left[1 - \left(\frac{1}{s} - \frac{1}{h}\right)\beta\right]^{-1}.$$

Simple substitution yields

$$T_D = T - T\beta\left(\frac{1}{i} + \frac{1}{s}\right). \tag{5}$$

Maximum profits are

$$\pi^* = \lambda w - \sqrt{2\lambda hA/\gamma}. \tag{6}$$

We note that this solution is similar, though not identical, to the inventory model with a stockout penalty. The critical difference is that this model allows increased sales not only at the beginning of each cycle, but also at time T_D: some customers anticipate the inability to purchase the good at the time most convenient to them and purchase the good before then.

We next analyze two extreme cases of the previous solution: those for which $T_D = 0$ and $T_D = T$. In the former solution, the firm holds no inventory and sells goods only at the instant it receives a new shipment. The latter solution is the EOQ one in which each buyer purchases the good at his w-time.

When $T_D = 0$, the firm holds no inventory, and makes sales only at intervals of T units of time. We wish to find the optimal values of p and T. According to Corollary 1, $T \leqslant (w - p)(1/i + 1/s)$. For a given value of T, profit maximization requires that p be as large as this constraint allows, so that $T = (w - p)(1/i + 1/s)$. Consequently, the firm's profits per unit of time are

$$\pi = -\frac{A}{T} + \lambda p = -\frac{A}{T} + \lambda\left(w - \frac{Tis}{i + s}\right), \tag{7}$$

which are maximized at

$$T = \sqrt{\left(\frac{1}{i} + \frac{1}{s}\right)\frac{A}{\lambda}} \tag{8}$$

and

$$p = w - \sqrt{\frac{(is)A}{\lambda(i + s)}}. \tag{9}$$

The profit per unit of time is

$$\pi^* = \lambda w - 2\sqrt{\frac{\lambda A i s}{i + s}}. \tag{10}$$

The other extreme case, with $T_D = T$ and $p = w$, is the EOQ model. The firm's objective is to maximize $\pi = -A/T + \lambda w - \lambda h T/2$; the optimal value of T statisfies

$$T = \sqrt{\frac{2A}{\lambda h}} \tag{11}$$

and the rate of profit is

$$\pi^* = \lambda w - \sqrt{2\lambda h A}. \tag{12}$$

For these values of T_D, T, and p, Equations 3–6 imply that $\beta = 0$ and that $\gamma = 1$. This last result is instructive. For the general case described by Equation 4, γ is less than or equal to one. Therefore, the length of a cycle under the EOQ solution is never smaller than the cycle length when the optimal solution calls for a no-sales interval.

Note that in the classic model with back-orders, inventory will *always* be held if the inventory carrying cost is finite and the stockout penalty is positive. In our model, however, it may happen that no inventory will be carried even for finite positive values of s, h, and i.

3. Optimal Policy when $s > i$

Consider next the optimal policy when s is greater than i. This section shows that the last two inventory policies discussed in Section 2 are the only ones that apply to this situation. In particular, if $h(1/i + 1/s) > 2$, sales will be made only at the instants at which the firm obtains new inventory. If $h(1/i + 1/s) < 2$, the firm will make sales throughout each cycle.

Proposition 3 states that when $s > i$, either T_D will equal T, or sales will be made at discrete points. The EOQ policy, for which $T_D = T$, has already been described in Section 2. The main concern in this discussion will be with a policy of discrete sales. That policy can take only one form:

Proposition 4. *Let s be greater than i, and let T be the length of an inventory cycle. Of all policies that involve sales at discrete points only, the most profitable one requires that sales be made only at* $0, T, 2T \ldots$

Proof. Suppose otherwise, that within some interval $[T - \tau, T]$, with $0 < \tau \leqslant T$, sales are made at times $T - \tau$, T and at some additional point, $T - \tau_1$. Suppose the other sales points in $[0, T - \tau]$ are fixed at their optimal values. We shall show that the optimal value of τ_1 is either 0 or τ; that is, for any value of τ, sales will be made only at the end points of the interval $[T - \tau, T]$.

Let K be the firm's inventory holding cost on goods sold to customers with w-times in $(0, T - \tau)$; any change in τ_1 will not affect this cost.

The firm's profits over a cycle are

$$\pi = -A - K + \lambda p T - \lambda h (T - \tau_1)\frac{\tau_1 s + (\tau - \tau_1) i}{i + s} - \lambda h (T - \tau)\frac{(\tau - \tau_1)s}{i + s}. \tag{13}$$

To determine the optimal value of τ_1, we must first determine the sign of $d^2p/d\tau_1^2$. Let $I_{\max}$ be the length of the longest no-sales interval in $[0, T - \tau]$. By Corollary 1, p is equal to either $w - I_{\max}(is)/(i + s)$, or to $w - \tau_1(is)/(i + s)$, or to $w - (\tau - \tau_1)(is)/(i + s)$, depending on which of $I_{\max}$, τ_1, and $\tau - \tau_1$ is the largest. In all cases, $d^2p/d\tau_1^2 = 0$ wherever the derivative exists. Thus $d^2\pi/d\tau_1^2 = 2\lambda h(s - i)/(i + s)$. This expression is positive, since, by assumption, $s > i$. Thus as a function of τ_1, π does not have an interior maximum in the interval $0 < \tau_1 \leqslant \tau$. The optimal value of τ_1 is either 0 or τ for *any* value of τ in $[0, T]$, and Proposition 4 follows.

Thus, any discrete policy requires sales only at times $0, T, 2T \ldots$ The optimal values of p and T and the profit per unit time are as in Equations 8–10. When $h(1/i + 1/s) < 2$, the optimal policy is EOQ and the optimal values of p and T are given in Equations 11 and 12.

4. Summary

We have seen that the optimal inventory policy can take only the following form: sell during the interval $[0, T_D]$, and do not sell during the interval (T_D, T), for a value of T_D equal to either 0, T, or some intermediate value. As shown in Figure 2, which of these three policies is optimal depends on the values of i, s, and h.

In Region *I*, $T_D = T$, and the inventory policy is EOQ. The condition for profitability is that $\lambda w - \sqrt{2\lambda h A} > 0$ (see Equation 12).

In Region II, $T_D = 0$: sales are made at intervals T units apart, coinciding with the firm's receipt of a shipment. The condition for profitability is that $\lambda w - 2\sqrt{\lambda A i s/(i + s)} > 0$ (see Equation 10).

In Region III, $0 < T_D < T$, and sales are continuously made during a portion of each cycle. The condition for profitability is that $\lambda w - \sqrt{2\lambda h A/\gamma} > 0$ (see Equation 6).

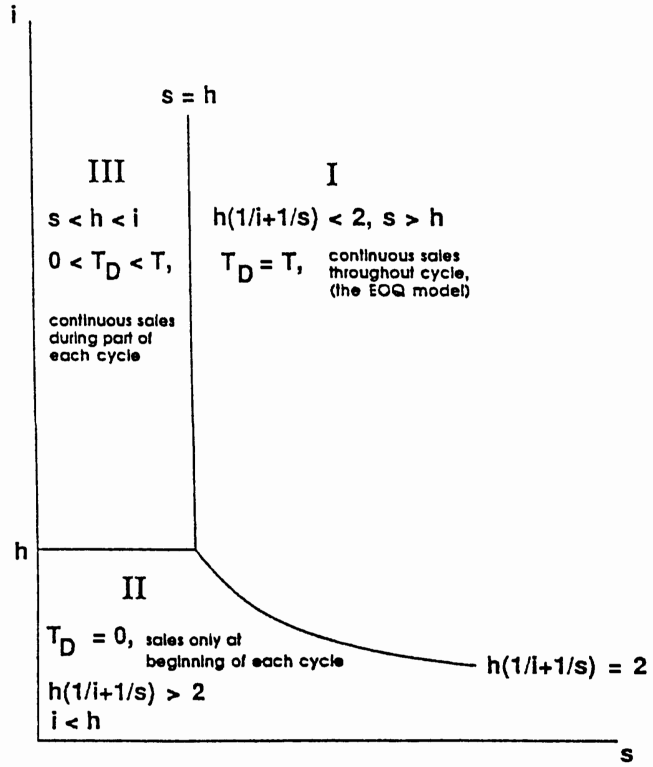

Figure 2. Optimal policies as a function of problem parameters.

If $s = h = i$, these policies are equally profitable. Indeed, in that case, a policy involving discrete sales within each cycle, and not only at its beginning, is optimal as well.

Our analysis is not exhaustive. We leave to future research an analysis that allows prices to vary over time, that considers customers who differ in ways other than their w-times, and that lets the rate of customer arrivals vary over time.

Acknowledgment

We are indebted to an associate editor and to an anonymous referee for several most helpful and insightful comments.

Dr. Hassin's research was supported by the Office of Naval Research, Contract ONR-N00014-79-C-0685 at the Center for Research on Organizational Efficiency, Institute of Mathematical Studies in the Social Sciences, Stanford University.

References

BUCOVETSKY, S. 1983. Price Dispersion and Stockpiling by Consumers. *Rev. Econ. Studies* **50,** 443–466.

HADLEY, G., AND T. M. WHITIN. 1963. *Analysis of Inventory Systems.* Prentice-Hall, Englewood Cliffs, N.J.

SOBEL, J. 1984. The Timing of Sales. *Rev. Econ. Studies* **51,** 353–368.