



University of Navarra

Working Paper
WP no 719
November, 2007

MYOPIC INVENTORY POLICIES USING INDIVIDUAL CUSTOMER ARRIVAL INFORMATION

Víctor Martínez de Albéniz¹

Alejandro Lago¹

¹ Professor, Operations Management and Technology, IESE

IESE Business School – University of Navarra

Avda. Pearson, 21 – 08034 Barcelona, Spain. Tel.: (+34) 93 253 42 00 Fax: (+34) 93 253 43 43

Camino del Cerro del Águila, 3 (Ctra. de Castilla, km 5,180) – 28023 Madrid, Spain. Tel.: (+34) 91 357 08 09 Fax: (+34) 91 357 29 13

Copyright © 2007 IESE Business School.

Myopic Inventory Policies Using Individual Customer Arrival Information¹

We investigate optimality of myopic policies using the single-unit decomposition approach in inventory management. We derive, under certain conditions, closed-form replenishment decisions, that we call a *base-probability policy*. That is, the order associated with a given customer is placed if and only if its arrival probability within the lead-time is higher than a threshold.

1. Introduction

Placing inventory buffers in the supply chain allows a better matching between supply and demand. The size of these buffers can be adjusted to provide an appropriate level of service to customers. In order to quantify replenishment decisions, traditional inventory models associate a cost to holding inventory and a back-ordering cost for making the customers wait for their orders. Holding costs account for the cost of working capital, invested in a product that has not been sold yet. Back-ordering costs, on the other hand, put a price to the waiting of the customer, who arrived before the product was available. By balancing these two costs appropriately, inventory is managed at the lowest cost.

When there are no set-up costs associated with an order, the optimal replenishment policy is often a *base-stock policy*: at each time period, there is an *optimal base-stock level*, and one should raise the current inventory level to that target level, or do nothing if the current level is already above the target. The inventory management literature is extensive on this point. The result is true in multi-echelon systems, see the seminal paper of Clark and Scarf [3]; for i.i.d. or correlated customer demands, see Chen and Song [2] or Song and Zipkin [10]; for fixed or random lead-times, see Kaplan [5]; and for non-stationary costs and prices, see Section 9.4.7 of Zipkin [13].

In general, closed-form solutions describing the base-stock level are available for simple situations, e.g., when lead-time is fixed, costs are stationary and demand is i.i.d. When the situation is non-stationary, very few formulas to compute the base-stock level analytically are available; one must often use numerical optimization or simulation. In particular, computing

¹An older version of this work was titled "Inventory Management by Synchronizing Replenishment Orders with Customers".

the optimal base-stock levels requires formulating a dynamic program (DP) that may suffer from the curse of dimensionality, as non-stationarity may expand the dimension of the state space in the DP.

Interestingly, some recent publications have incorporated new proof methods that avoid the dimensionality problem. They generalize the optimality of base-stock policies to different situations, see Axsater [1] and Muharremoglu and Tsitsiklis [8] and [9]. The crucial observation is that one can match each order placed by the inventory manager with a given customer. For example, the 5-th order will fulfill the 5-th unit of demand. Using this matching, as shown in Muharremoglu and Tsitsiklis [8], one can decouple the ordering decision unit by unit, and decide whether a unit should be ordered *independently of all other units*. This approach allows to define several simpler dynamic program with a smaller state space.

In this paper, we build on the approach of Muharremoglu and Tsitsiklis [8], in a context of single echelon, uncertain demand, cost and price, and fixed lead-time. While their paper focuses on showing the optimality of base-stock policies, we concentrate on operationalizing the ordering policies, by providing, under certain conditions, closed-form formulas to determine whether to order or not.

Specifically, within the single-unit decomposition approach, we provide conditions under which a myopic policy is optimal. Some of these conditions are related to the ones provided in the early papers of Veinott [11] and [12], and Lovejoy [7]. However, our condition on the demand process is more general than what is usually assumed: Karlin [6], Veinott [12] or Song and Zipkin [10], for example, require that the demand is stochastically increasing, while we only require that the arrival probability of a certain customer increases over time. Furthermore, we develop a simple analytical formula to decide whether to place an order or not: for each specific customer, an order should be placed if and only if its probability of arrival within the lead-time is high enough. While this is theoretically equivalent to an optimal base-stock level, conceptually it allows the replenishment decision to be taken customer by customer.

The paper starts with the description of the model in Section 2. Section 3 develops the main results. Finally, we conclude the paper with a discussion in section 4. All the proofs can be found in the Appendix.

2. The Model

Consider a firm that distributes a single product to customers, in an infinite horizon setting. This product is procured from an external supplier who is located far away, and takes L time units to deliver an order to the firm. Lead-time is fixed, but the methodology could be used in a similar way for stochastic lead-times, as soon as the ordering sequence and the receiving sequence are identical, i.e., orders do not cross, see Kaplan [5] and Muharremoglu and Tsitsiklis [9]. The inventory is managed using a standard periodic-review system with back-ordering. At each time period $t = 1, \dots, \infty$, the firm first checks the inventory level, places an order q_t to the supplier, which will be received at time $t + L$. Then customers arrive, and are served if there is stock on-hand; otherwise, they are left on a waiting line, and will be served first-come-first-served when more inventory arrives. Finally, at the end of the period, a per-unit inventory holding fee h (fixed and unrelated to the purchasing cost, since the capital cost of inventory is taken into consideration by a discount factor) and a per-unit backlogging penalty b (also fixed) are charged.

At each period t , all the information on past and present costs, prices and demands is available, and denoted I_t . Based on this, the firm can generate a distribution on future events. We denote by $\mathbb{P}_{I_t}(A)$ the probability of event A conditional on the present information I_t . Similarly, $\mathbb{E}_{I_t}(X)$ denotes the expectation of a random variable X conditional on the information I_t .

At each time period t , a stochastic number of customers D_t arrives. Its distribution is completely determined by I_t . In addition, these customers come in a given sequence. We denote by T_k the arrival time of the k -th customer. It is clear that all distributional information about the demand process can be translated into the arrival process, since for all k, t' ,

$$\mathbb{P}_{I_t} \left(\sum_{\tau=1}^{t'} D_\tau \geq k \right) = \mathbb{P}_{I_t} (T_k \leq t'). \quad (1)$$

The per-unit purchasing cost charged by the supplier is denoted by C_t , and can change stochastically from period to period. Its evolution depend exclusively on I_t . Thus, the expected cost for period $t + \Delta t$, at time t , is denoted by $\mathbb{E}_{I_t} C_{t+\Delta t}$.

Similarly, there is a per-unit selling price of P_t , that can also change stochastically over time. When a customer that arrives at t can be served immediately, the firm receives P_t immediately. However, if there is no inventory available on-hand, and the customer is served

at t' , the firm receives only $r(P_t, P_{t'})$, at the delivery time. This is a flexible approach to prices, that allows to charge the price on the arrival time, i.e., $r(P_t, P_{t'}) = P_t$, or the price on the delivery time, i.e., $r(P_t, P_{t'}) = P_{t'}$.

Finally, we consider a discount rate of α across periods, corresponding to the time-value of money. As it is common in the inventory management literature, we assume that the firm is risk-neutral. The objective is to maximize the net present value (NPV) of the firm, also called discounted profit-to-go, by selecting the most appropriate inventory policy, i.e., the ordering time t_k of the k -th order,

$$\begin{aligned} \max_{t_1, t_2, \dots} \mathbb{E}_{I_0} & \left\{ \sum_{t=1}^{\infty} \alpha^t \left[\begin{array}{c} h \left(\sum_{k=1}^{\infty} 1_{t_k+L \leq t \leq T_k} \right) + b \left(\sum_{k=1}^{\infty} 1_{T_k \leq t \leq t_k+L} \right) \\ -C_t \left(\sum_{k=1}^{\infty} 1_{t=t_k} \right) + \sum_{k=1}^{\infty} r \left(P_{T_k}, P_{\max\{T_k, t_k+L\}} \right) 1_{t=\max\{T_k, t_k+L\}} \end{array} \right] \right\} \\ & = \sum_{k=1}^{\infty} \mathbb{E}_{I_0} \left\{ \sum_{t=1}^{\infty} \alpha^t \left[\begin{array}{c} h 1_{t_k+L \leq t \leq T_k} + b 1_{T_k \leq t \leq t_k+L} \\ -C_t 1_{t=t_k} + r \left(P_{T_k}, P_{\max\{T_k, t_k+L\}} \right) 1_{t=\max\{T_k, t_k+L\}} \end{array} \right] \right\}, \end{aligned} \quad (2)$$

where $1_A = 1$ if A is true and 0 otherwise. Here, we can use the decomposition approach of Muharremoglu and Tsitsiklis [8]. This is possible since by lead-time is fixed and the backlogging assumption guarantees that the demand process is independent of the order process. Thus, the maximization problem of Equation (2) can be decomposed in an independent problem for each k , as follows,

$$\mathbb{E}_{I_0} \left\{ h \left(\sum_{t=t_k+L}^{T_k} \alpha^t \right) + b \left(\sum_{t=T_k}^{t_k+L} \alpha^t \right) - C_{t_k} \alpha^{t_k} + r \left(P_{T_k}, P_{\max\{T_k, t_k+L\}} \right) \alpha^{\max\{T_k, t_k+L\}} \right\}. \quad (3)$$

Assuming that the first $k-1$ orders have been placed, and that we need to place the k -th, consider the decision at time period t . The decision tree for ordering the k -th unit can be summarized in what follows.

1. Place the order now, at period t . In that case, the discounted profit-to-go of this decision is

$$U(k, t, I_t) := \mathbb{E}_{I_t} \left\{ h \left(\sum_{\tau=t+L}^{T_k} \alpha^{\tau} \right) + b \left(\sum_{\tau=T_k}^{t+L} \alpha^{\tau} \right) - C_t \alpha^t + r \left(P_{T_k}, P_{\max\{T_k, t+L\}} \right) \alpha^{\max\{T_k, t+L\}} \right\}.$$

2. Or wait and see; in that case, we obtain updated information on the demand, price and cost processes, and we face the same decision (order or not) at period $t+1$. That

is, we can either place the order at $t + 1$, with discounted profit

$$\mathbb{E}_{I_{t+1}} \left\{ h \left(\sum_{\tau=t+L+1}^{T_k} \alpha^\tau \right) + b \left(\sum_{\tau=T_k}^{t+L+1} \alpha^\tau \right) - C_{t+1} \alpha^{t+1} + r \left(P_{T_k}, P_{\max\{T_k, t+L+1\}} \right) \alpha^{\max\{T_k, t+L+1\}} \right\},$$

or wait and see (and go into the next branch in the tree).

We see that purchasing an item at t amounts to comparing the profit-to-go of this decision, denoted $U(k, t, I_t)$, with the profit-to-go of delaying the purchase. Let $V(k, t, I_t)$ be the value of purchasing the k -th item at t or later, with the information at time t . Hence, the optimization program can be expressed as the following dynamic program, solved by backwards recursion:

$$V(k, t, I_t) = \max\{U(k, t, I_t), \mathbb{E}_{I_t} V(k, t + 1, I_{t+1})\}. \quad (4)$$

Note that in the standard inventory management approach, the state space includes the inventory level and the demand forecast for all future periods, i.e., we must consider the probability that $D_\tau = d$ for each d and for each period τ . In our model, we decompose the problem for each unit k , and thus we only require the forecast distribution of T_k . Of course, we need to compute a DP for each different k . In addition, this approach allows us to obtain analytical formulas for each k , as shown in the next section.

3. Optimality of Myopic Policies

Under a number of assumptions, we can characterize $V(k, t, I_t)$ in a simple way. These assumptions allow to simplify the dynamic program so that a myopic, one-step look-ahead, policy is optimal, and thus a closed-form formula is available. In the literature, see Veinott [11] or [12] for example, a myopic policy is shown to be optimal when the demand is stochastically increasing and some monotonicity requirements are placed on the cost and price processes.

Our regularity assumptions are similar. First, we require the demand process to exhibit a monotonicity property, which is weaker than being stochastically increasing in time: we assume that the arrival time of each customer minus the current time is stochastically non-increasing. Second, we need the price and cost processes to satisfy a monotonicity property, similar to the literature.

We start with a preliminary lemma.

Lemma 1 Consider for all k , $1\{U(k, t, I_t) - \mathbb{E}_{I_t}U(k, t + 1, I_{t+1}) \geq 0\}$, i.e., 1 if the event occurs, and 0 otherwise, and assume that it is stochastically non-decreasing in t (in each sample path). Then a myopic policy is optimal, i.e., $U(k, t, I_t) = V(k, t, I_t)$ if and only if $U(k, t, I_t) \geq \mathbb{E}_{I_t}U(k, t + 1, I_{t+1})$.

This lemma provides a sufficient condition for myopic policies to be optimal. To obtain the desired condition for Lemma 1, we focus on the following class of demand processes.

Assumption 1 Any customer gets closer when time advances. For all k , for all Δt , for each sample path,

$$\mathbb{P}_{I_t}(T_k \leq t + \Delta t) \leq \mathbb{P}_{I_{t+1}}(T_k \leq t + \Delta t + 1). \quad (5)$$

That is, the chances of customer k arriving before Δt units gets larger as time advances, regardless of the information acquired between t and $t + 1$.

The interpretation is the following. Consider at t , with all the available information, the probability that the k -th customer arrives within Δt periods. Then, when one incorporates the information update at $t + 1$, the probability of arrival within the same Δt periods must go up, regardless of the information update. That is, the customer's likelihood of arrival can never decrease. This assumption is weaker than having stochastically increasing demands, used in Veinott [12], Karlin [6] or Song and Zipkin [10]. Indeed, consider that the demand arriving per period is independent over time, and stochastically increasing. Without loss of generality, it is sufficient to analyze $t = 1$ and k to show that Assumption 1 holds. For any $d_1 \geq 0$,

$$\begin{aligned} \mathbb{P}_1(T_k \leq \tau) &= \mathbb{P}_1(D_1 + \dots + D_\tau \geq k) \\ &\leq \mathbb{P}_1(D_2 + \dots + D_{\tau+1} \geq k) \text{ since } D_{\tau+1} \succeq D_1 \\ &\leq \mathbb{P}_2(D_2 + \dots + D_{\tau+1} \geq k - d_1) \\ &= \mathbb{P}_2(T_k \leq \tau + 1 | D_1 = d_1). \end{aligned}$$

In addition, it contains demand processes that are not stochastically increasing. For example, when the demand process is generated by customer arrivals with exponential inter-arrival times of decreasing rate as the customer rank increases, then Assumption 1 is satisfied (see example below), but the demand is stochastically non-increasing.

Some demand processes do not satisfy the assumption, such many ARMA processes. Interestingly, these instances, in the case of ARMA, could be modified so that they fit the assumption, see Johnson and Thompson [4]. By assuming some minimum level of demand,

one is able to guarantee that the realized inventory levels are always below the myopic base-stock levels, yielding optimality of myopic policies. Our assumption provides a similar effect.

Also, the condition may be violated for heavy-tailed inter-arrival times, but is always satisfied when the inter-arrival times have a non-decreasing failure rate, as show in the next example.

Example 1 Non-decreasing failure rates. *Assume that the demand is generated by a queueing model of arrivals of consecutive customers. Assumption 1 is satisfied when inter-arrival times are i.i.d. with a non-decreasing failure rate, i.e., $\mathbb{P}(T = t|T \geq t)$ non-decreasing. This holds for Poisson arrivals. Also, if \bar{p}_t is the probability that the inter-arrival time is t or larger, then the condition is satisfied when*

$$\frac{\bar{p}_t - \bar{p}_{t+1}}{\bar{p}_t} \leq \frac{\bar{p}_{t+1} - \bar{p}_{t+2}}{\bar{p}_{t+1}},$$

which is equivalent to $\frac{\bar{p}_t}{\bar{p}_{t+1}}$ non-decreasing.

Furthermore, we can show that neither this condition nor the assumption is satisfied for heavy-tailed inter-arrival distributions, i.e., when the decay of \bar{p}_t is slower than any exponential, e.g. when $\bar{p}_t = \frac{1}{1+t}$.

We use a second assumption to simplify the analysis. This is commonly assumed in the inventory management literature.

Assumption 2 Price and demand processes are independent.

Under these assumptions, we can prove the following theorems.

Theorem 1 *Assume that the price is determined when the order is made, i.e., $r(P_{arriv}, P_{deliv}) = P_{arriv}$. If Assumptions 1 and 2 hold, if $C_t - \alpha \mathbb{E}_{I_t} C_{t+1}$ is non-increasing for each sample path, if for all $\tau \in [0, \dots, L-1]$ $\mathbb{E}_{I_t} (P_{t+\tau} - P_{t+\tau+1})$ is non-decreasing for each sample path, and if $\mathbb{E}_{I_t} P_{t+L}$ is non-decreasing for each sample path, then the following is true:*

- (i) *If it is optimal to order at t , it is also optimal to order at $t+1$ regardless of the information received between t and $t+1$.*

(ii) *Base-probability policy: the k -th order must be placed at time t if and only*

$$b + C_t - \alpha \mathbb{E}_{I_t} C_{t+1} \leq \sum_{\tau=0}^{L-1} \alpha^L (1 - \alpha) \mathbb{E}_{I_t} (P_{t+\tau} - P_{t+\tau+1}) \mathbb{P}_{I_t}(T_k \leq t + \tau) + \left(b + h + \alpha^L (1 - \alpha) \mathbb{E}_{I_t} P_{t+L} \right) \mathbb{P}_{I_t}(T_k \leq t + L). \quad (6)$$

Theorem 2 *Assume that the price is determined when the order is delivered, i.e., $r(P_{arriv}, P_{deliv}) = P_{deliv}$. If Assumptions 1 and 2 hold, and if $\frac{b + C_t - \alpha \mathbb{E}_{I_t} C_{t+1}}{b + h + \alpha^L \mathbb{E}_{I_t} (P_{t+L} - \alpha P_{t+L+1})}$ is non-increasing for each sample path, then the following is true:*

(i) *If it is optimal to order at t , it is also optimal to order at $t + 1$ regardless of the information received between t and $t + 1$.*

(ii) *Base-probability policy: the k -th order must be placed at time t if and only if*

$$\frac{b + C_t - \alpha \mathbb{E}_{I_t} C_{t+1}}{b + h + \alpha^L \mathbb{E}_{I_t} (P_{t+L} - \alpha P_{t+L+1})} \leq \mathbb{P}_{I_t}(T_k \leq t + L). \quad (7)$$

The theorems provide a closed-form condition for the replenishment decision. In fact, Equations (6) and (7) are equivalent to $U(k, t, I_t) - \mathbb{E}_{I_t} U(k, t + 1, I_{t+1}) \geq 0$. The meaning of these equations is intuitive: when the k -th customer is getting close, measured by the probability of arriving within a given number of periods, the order must be placed. We call this a *base-probability policy* since the order is placed only when the arrival probability within the lead-time is higher than a threshold. Note that this corresponds to a state-dependent base-stock policy in traditional inventory management models.

Notice that the result can be easily extended to continuous time, where information updates over T_k flow continuously. It is interesting to note that, in Theorem 2, the optimal policy comes from comparing a term that depends on the cost and price processes with a term that depends on the demand process.

The assumptions on the price and cost processes required in the theorems are satisfied for many simple situations. Of course, they are true when C_t and P_t are deterministic and stationary. In that case, both theorems provide the same order condition:

$$\frac{b + (1 - \alpha)C}{b + h + \alpha^L (1 - \alpha)P} \leq \mathbb{P}_t(T_k \leq t + L).$$

One can also consider the case where $P_t = p$ and C_t is stochastic such that $C_{t+1} = C_t(1 - \epsilon_t)$ where the cost decreases by $\epsilon_t \geq 0$, and has a stationary average $\mathbb{E}_{I_t} \epsilon_t = \mu$. Other instances

include, in the case of $r(P_{arriv}, P_{deliv}) = P_{deliv}$, situations where the price process is equal to the cost process plus a fixed mark-up, i.e., $P_t = C_t + m$, and $C_{t+1} = C_t(1 - \epsilon_t)$, defined as before.

When the conditions of the theorem are not satisfied, the myopic policy may not be optimal, and one should resort to a numerical method to solve the dynamic program, i.e., Equation (4). We show next two examples where the myopic policy is not optimal, in the case of prices being determined at delivery. In each case, one of the two assumptions is not satisfied.

Example 2 Heavy-tailed inter-arrival times. *Assume that price and cost are stationary, equal to p and c respectively, that $h = b = 0$ and that $r(P_{arriv}, P_{deliv}) = P_{deliv}$. Thus, the left-hand side of (7) is constant. Consider $k = 1$ and that the arrival time of the (first) customer is $t \geq 1$ with probability $\frac{1}{t} - \frac{1}{t+1} = \frac{1}{t(t+1)}$, that is, heavy-tailed distributed and hence, not satisfying Assumption 1.*

We can show (details in the appendix) that the myopic policy is not optimal. Indeed, with this type of demand when the customer arrives late, it tends to arrive very late. The myopic policy underestimates the value of the information update and thus, suggests to place the order earlier than it should.

An numerical illustration is provided in Figure 1 (left). In the figure, the myopic policy dictates that one should place the order for $t \leq 3$, that is when $U(t, \text{not arrived}) \geq \mathbb{E}_{I_t=\text{not arrived}}U(t+1)$. However, since $V(t, \text{not arrived}) > U(t, \text{not arrived})$ for all t , it is never optimal to place an order when the customer has not arrived.

As a consequence, the results of Theorem 2 cannot hold when we remove Assumption 1.

Example 3 Increasing costs. *Assume that the arrival time of the first customer is exponential, i.e., the probability of arrival on $t+1$ given that it has not arrived at t is $1-\beta \in (0, 1)$. Consider now a stationary price p but a cost $C_t = p\alpha^L(1 - \theta^t)$ that increases over time. Let $h = b = 0$. Thus, the discounted margin $p - \alpha^L C_t$ decreases by a factor $\theta < 1$ per period. Also, $C_t - \alpha \mathbb{E}_{I_t} C_{t+1}$ increases. Assumption 1 is satisfied, but the left-hand side of (7) is increasing.*

We show (details in the appendix) that the myopic policy is not optimal. The intuition is that sometimes, it may happen $0 \geq U(t, \text{not arrived}) \geq \mathbb{E}_{I_t=\text{not arrived}}U(t+1)$. The myopic policy may suggest to place an order even though it is not profitable to do so on expectation, because it focuses on the potential margin loss of delaying the sale, and neglects the value of

acting only when the customer has arrived. Thus, it underestimates the value of delaying the ordering decision.

A numerical example is provided in Figure 1 (right). The figure indicates that it is optimal to place an order for $t \leq 7$, while the myopic policy yields placing the order for $t \leq 24$. Hence, the results of Theorem 2 do not necessarily hold when we remove that the left-hand side of (7) is non-increasing.

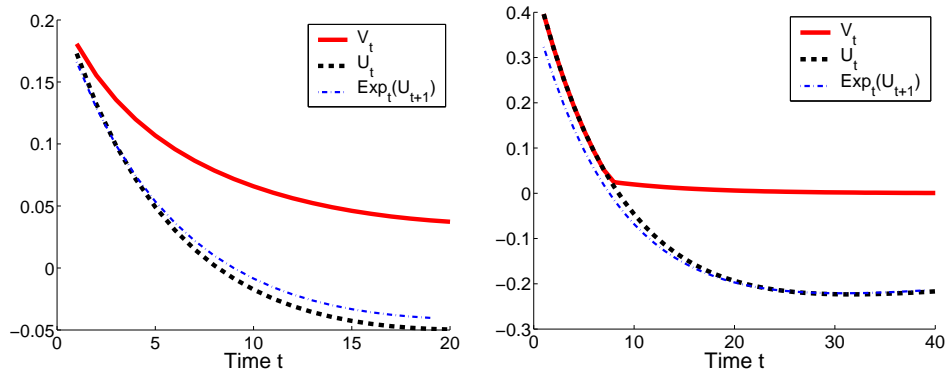


Figure 1: Plot of $V(t, \text{not arrived})$, $U(t, \text{not arrived})$, and $\mathbb{E}_{(I_t=\text{not arrived})}U(t+1)$ for Examples 2 (left) and 3 (right). On the left figure, the parameters are $L = 5, p = 1, c = 0.5, h = b = 0$ and $\alpha = 0.95$. On the right figure, $L = 20, p = 1, h = b = 0, \alpha = 0.99, \beta = 0.99$ and $\theta = 0.9$.

4. Conclusion

The model presented in this paper uses the single-unit decomposition framework to derive optimality of myopic policies under certain conditions. These conditions, specifically those on the demand process, are weaker than having stochastically increasing demands across time. Our approach yields a closed-form order policy, what we call a *base-probability policy*. This policy dictates that the order of customer k should be placed at t if and only if the customer arrival probability within the lead-time is higher than a certain threshold determined by the cost and price processes.

The methodology applied in the paper can be extended directly to batch ordering. Other more general situations, such as the stochastic lead-time case with non-crossing orders, can also be approached with the same method but the resulting ordering rules are not as simple in this case.

Acknowledgements

We would like to thank the associate editor and three anonymous referees for helping us improve significantly this manuscript.

References

- [1] Axsater S. 1990. "Simple Solution Procedures for a Class of Two-Echelon Inventory Problems." *Operations Research*, **38**(1), pp. 64-69.
- [2] Chen F. and J.-S. Song 2001. "Optimal Policies For Multiechelon Inventory Problems With Markov-Modulated Demand." *Operations Research*, **49**(2), pp. 226-234.
- [3] Clark A. J. and H. Scarf 1960. "Optimal Policies for a Multi-Echelon Inventory Problem." *Management Science*, **6** (4), pp. 475-490.
- [4] Johnson G. D. and H. E. Thompson 1975. "Optimality of Myopic Inventory Policies for Certain Dependent Demand Processes." *Management Science*, **21**(11), pp. 1303-1307.
- [5] Kaplan R. S. 1970. "A Dynamic Inventory Model with Stochastic Lead Times" *Management Science*, **16**(7), pp. 491-507.
- [6] Karlin S. 1960. "Dynamic Inventory Policy with Varying Stochastic Demands." *Management Science*, **6**(3), pp. 231-258.
- [7] Lovejoy W. S. 1992. "Stopped Myopic Policies in Some Inventory Models with Generalized Demand Processes." *Management Science*, **38**(5), pp. 688-707.
- [8] Muharremoglu A. and J. N. Tsitsiklis 2001. "A Single-Unit Decomposition Approach to Multi-Echelon Inventory Systems." Working paper, Graduate School of Business, Columbia University.
- [9] Muharremoglu A. and J. N. Tsitsiklis 2003. "Dynamic Leadtime Management in Supply Chains." Working paper, Graduate School of Business, Columbia University.
- [10] Song J. S. and P. H. Zipkin 1993. "Inventory Control in a Fluctuating Demand Environment." *Operations Research*, **41**(2), pp. 351-370.
- [11] Veinott A.F. Jr. 1965. "Optimal Policy in a Dynamic, Single Product, Nonstationary Inventory Model with Several Demand Classes." *Operations Research*, **13**(5), pp. 761-778.
- [12] Veinott A.F. Jr. 1965. "Optimal Policy for a Multi-Product, Dynamic, Nonstationary Inventory Problem." *Management Science*, **12**(3), pp. 206-222.
- [13] Zipkin P. H. 2000. *Foundations of Inventory Management*. McGraw-Hill International Editions.

Proof of Lemma 1

Proof. If $U(t, k, I_t) - \mathbb{E}_{I_t}U(k, t + 1, I_{t+1}) \geq 0$ is non-decreasing for all sample paths, then if $U(t, k, I_t) - \mathbb{E}_{I_t}U(k, t + 1, I_{t+1}) \geq 0$, then the same is true for $t + 1$, i.e., $U(t + 1, k, I_{t+1}) - \mathbb{E}_{I_{t+1}}U(k, t + 2, I_{t+2}) \geq 0$, and so on. A value iteration argument yields that $V(k, t, I_t) = U(k, t, I_t)$. On the other hand, if $U(t, k, I_t) - \mathbb{E}_{I_t}U(k, t + 1, I_{t+1}) < 0$, then $V(k, t, I_t) \geq \mathbb{E}_{I_t}U(k, t + 1, I_{t+1}) > U(t, k, I_t)$. ■

Proof of Theorems 1 and 2

Proof. We can calculate

$$U(k, t, I_t) - \mathbb{E}_{I_t}U(k, t + 1, I_{t+1}) = \alpha^t \left(\begin{array}{l} -b + (h + b)\mathbb{P}_{I_t}(T_k \leq t + L) - C_t + \alpha\mathbb{E}_{I_t}C_{t+1} \\ + \alpha^L \mathbb{E}_{I_t} \left\{ 1_{t+L \geq T_k} \left(r(P_{T_k}, P_{t+L}) - \alpha r(P_{T_k}, P_{t+L+1}) \right) \right\} \end{array} \right)$$

When the price is paid when the order is made, then

$$U(k, t, I_t) - \mathbb{E}_{I_t}U(k, t + 1, I_{t+1}) = \alpha^t \left(\begin{array}{l} -b - C_t + \alpha\mathbb{E}_{I_t}C_{t+1} \\ + [\alpha^L(1 - \alpha)\mathbb{E}_{I_t} \{P_{T_k} | T_k \leq t + L\} + (h + b)] \mathbb{P}_{I_t}(T_k \leq t + L) \end{array} \right)$$

and when it is paid when the order is delivered, then

$$U(k, t, I_t) - \mathbb{E}_{I_t}U(k, t + 1, I_{t+1}) = \alpha^t \left(\begin{array}{l} -b - C_t + \alpha\mathbb{E}_{I_t}C_{t+1} \\ + [\alpha^L \mathbb{E}_{I_t} \{P_{t+L} - \alpha P_{t+L+1} | T_k \leq t + L\} + (h + b)] \mathbb{P}_{I_t}(T_k \leq t + L) \end{array} \right)$$

We simply apply the assumptions to show that if Equations (6) and (7) are satisfied at t , they are also satisfied at $t + 1$, for each sample path. Lemma 1 yields the theorems. ■

Details of Example 2

If the customer has still not arrived yet at time t , the conditional probability that the customer arrives at $t + \Delta t \geq t + 1$ is $\gamma_{t, t+\Delta t} = \frac{t + 1}{(t + \Delta t)(t + \Delta t + 1)}$. Hence,

$$U(t, \text{arrived}) = \alpha^t(p\alpha^L - c)$$

$$U(t, \text{not arrived}) = \alpha^t \left\{ p \left(\frac{L}{t + L + 1} \alpha^L + \sum_{k=L+1}^{\infty} \frac{t + 1}{(t + k)(t + k + 1)} \alpha^k \right) - c \right\} < U(t, \text{arrived})$$

Thus,

$$\mathbb{E}_{I_t=\text{not arrived}} U(t+1) = \alpha^t \left\{ p \left(\frac{L+1}{t+L+2} \alpha^{L+1} + \sum_{k=L+2}^{\infty} \frac{t+1}{(t+k)(t+k+1)} \alpha^k \right) - c \alpha \right\}$$

Hence, $U(t, \text{not arrived}) \geq \mathbb{E}_{I_t=\text{not arrived}} U(t+1)$ if and only if

$$\frac{L + \alpha(L+2)}{t+L+1} - \frac{\alpha(L+1) + \alpha(L+2)}{t+L+2} \geq \frac{(1-\alpha)c}{\alpha^L p}.$$

If $L \leq \alpha(L+1)$, the left-hand side is decreasing. Otherwise, the left-hand side can be shown to be decreasing and then increasing to zero. Thus, in the general case, the condition is satisfied for $t \leq t_1$, where t_1 is the unique equation to

$$\frac{L + \alpha(L+2)}{t_1 + L + 1} - \frac{\alpha(L+1) + \alpha(L+2)}{t_1 + L + 2} = \frac{(1-\alpha)c}{\alpha^L p} = r.$$

That is,

$$t_1 = \frac{1}{2} \sqrt{\left(\frac{\alpha(L+1) - L}{r} \right)^2 + 4 \frac{L + \alpha(L+2) - r}{r}} - \frac{1}{2} \frac{\alpha(L+1) - L}{r}.$$

Thus, the condition can only be satisfied when $r = \frac{(1-\alpha)c}{\alpha^L p} \leq L + \alpha(L+2)$.

For large t , $U(t, \text{not arrived}) \leq 0$, and therefore it is optimal to produce only upon arrival.

Thus, there is t_2 such that for $t \geq t_2$, we can show that

$$\begin{aligned} V(t, \text{arrived}) &= \alpha^t (p\alpha^L - c) \\ V(t, \text{not arrived}) &= \alpha^t \left\{ \sum_{k=1}^{\infty} \frac{t+1}{(t+k)(t+k+1)} \alpha^k (p\alpha^L - c) \right\}. \end{aligned}$$

At $t = t_2 - 1$ we have that $U(t, \text{not arrived}) \geq \mathbb{E}_{I_t=\text{not arrived}} U(t+1)$, which is equivalent to

$$\begin{aligned} & p \left(\frac{L}{t+L+1} \alpha^L + \sum_{k=L+1}^{\infty} \frac{t+1}{(t+k)(t+k+1)} \alpha^k \right) - c \\ & \geq \sum_{k=1}^{\infty} \frac{t+1}{(t+k)(t+k+1)} \alpha^k (p\alpha^L - c), \end{aligned}$$

This is equivalent to

$$\begin{aligned} & \sum_{k=L+1}^{\infty} \frac{t+1}{(t+k)(t+k+1)} (\alpha^L - \alpha^k) \\ & \leq \left(\frac{L}{t+L+1} - \sum_{k=1}^L \frac{t+1}{(t+k)(t+k+1)} \alpha^k \right) \left(\alpha^L - \frac{c}{p} \right). \end{aligned}$$

The conclusion is straightforward: the myopic policy cannot be optimal.

Details of Example 3

$$\begin{aligned}
U(t, \text{arrived}) &= \alpha^t(p\alpha^L - C_t) = p\alpha^L\alpha^t\theta^t \\
U(t, \text{not arrived}) &= \alpha^t \left\{ p \left(\alpha^L(1 - \beta^L) + \frac{\alpha^{L+1}\beta^L(1 - \beta)}{1 - \alpha\beta} \right) - p\alpha^L(1 - \theta^t) \right\} \\
&= p\alpha^L\alpha^t \left(\theta^t - \beta^L \frac{1 - \alpha}{1 - \alpha\beta} \right) \\
&< U(t, \text{arrived}) \\
\mathbb{E}_{I_t=\text{not arrived}} U(t+1) &= p\alpha^L\alpha^{t+1} \left(\theta^{t+1} - \beta^{L+1} \frac{1 - \alpha}{1 - \alpha\beta} \right)
\end{aligned}$$

Hence, $U(t, \text{not arrived}) \geq \mathbb{E}_{I_t=\text{not arrived}} U(t+1)$ if and only if $\theta^t \geq \frac{\beta^L(1 - \alpha)}{1 - \alpha\theta}$, that is, $t \leq t_1$ for t_1 defined appropriately.

In addition, as in the previous example, for large t it is not profitable to place the order before the customer has arrived, since the expected profit from doing so is negative. This implies that for t large enough,

$$\begin{aligned}
V(t, \text{not arrived}) &= \sum_{k=1}^{\infty} \alpha^{t+k}(1 - \beta)\beta^{k-1}p\alpha^L\theta^{t+k} \\
&= p\alpha^L\alpha^t\theta^t \left(\frac{(1 - \beta)\alpha\theta}{1 - \alpha\beta\theta} \right)
\end{aligned}$$

Hence, at the last period t where the order is launched before arrival, we have that

$$\theta^t - \beta^L \frac{1 - \alpha}{1 - \alpha\beta} \geq \theta^t \left(\frac{(1 - \beta)\alpha\theta}{1 - \alpha\beta\theta} \right),$$

or equivalently,

$$\theta^t \left(\frac{1 - \alpha\theta}{1 - \alpha\beta\theta} \right) \geq \frac{\beta^L(1 - \alpha)}{1 - \alpha\beta}.$$

Hence, the myopic policy cannot be optimal.