A Single-Item Inventory Model for a Nonstationary Demand Process

Stephen C. Graves
Leaders for Manufacturing Program and A. P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge Massachusetts 02139-4307

In this paper, we consider an adaptive base-stock policy for a single-item inventory system, where the demand process is nonstationary. In particular, the demand process is an integrated moving average process of order (0, 1, 1), for which an exponential-weighted moving average provides the optimal forecast. For the assumed control policy we characterize the inventory random variable and use this to find the safety stock requirements for the system. From this characterization, we see that the required inventory, both in absolute terms and as it depends on the replenishment lead-time, behaves much differently for this case of nonstationary demand compared with stationary demand. We then show how the single-item model extends to a multi-stage, or supply-chain context; in particular we see that the demand process for the upstream stage is not only nonstationary but also more variable than that for the downstream stage. We also show that for this model there is no value from letting the upstream stages see the exogenous demand. The paper concludes with some observations about the practical implications of this work.
(*Single-Item Inventory Model; Nonstationary Demand; Base-Stock Policy; Amplification of Demand Variability Across a Supply Chain*)

1. Introduction

One major theme in the continuing development of inventory theory is to incorporate more realistic assumptions about product demand into inventory models. In most industrial contexts, demand is uncertain and hard to forecast. Many demand histories behave like random walks that evolve over time with frequent changes in their directions and rates of growth or decline. Furthermore, as product life cycles get shorter, the randomness and unpredictability of these demand processes have become even greater.

In practice, for such demand processes, inventory managers often rely on forecasts based on a time series of prior demand, such as a weighted moving average. Typically these forecasts are predicated on a belief that the most recent demand observations are the best predictors for future demand. A forecast based on an exponential-weighted moving average is a good and common example of this practice.

In this paper we develop a simple inventory model that incorporates this type of demand process, namely a demand process that behaves like a random walk. We consider a class of nonstationary demand processes, for which an exponential-weighted moving average provides the minimum mean square forecast; we then build a single-item inventory model assuming a deterministic replenishment lead-time for this family of demand. From the analysis of this model, we determine the safety stock requirements for a single item, and then explore implications from applying the model to a multi-stage or supply chain setting.

This work is related to a series of papers that develop optimal inventory policies when the demand distribution depends upon some unknown parameter

1523-4614/99/0101/0050$05.00

and an estimate of the parameter is updated as actual demand is observed over time. For demand distributions from the exponential family, Scarf (1959) formulates a single-product model as a dynamic program with a two-variable state space, and establishes that a base-stock policy based on a critical fractile is optimal. Scarf (1960) shows how to reformulate this problem as a single-variable dynamic program, subject to some additional assumptions. Azoury (1985) and Miller (1986) generalize and extend these results to other classes of demand distributions. Lovejoy (1990) shows that a critical-fractile inventory policy is optimal or near optimal for a more general class of demand distributions.

Veinott (1965) introduces the idea of a myopic solution in which the single-variable dynamic program can be solved with only information from the current period. He establishes the optimality of base-stock policies when demand is independent over time and there is a constant replenishment lead-time; he also characterizes conditions under which the myopic policy is optimal when demand is correlated. Johnson and Thompson (1975) extend the results of Veinott. They consider a single item with an autoregressive, moving-average (ARMA) demand process, zero replenishment lead-time and no backorders. They establish the optimality of a myopic policy under the assumption that there is a minimum and maximum bound on demand in each period, which imposes additional conditions on the parameters for the demand process.

Song and Zipkin (1993) present another single-item model with nonstationary demand. They assume that the demand process is Poisson, but where a Markov process governs the rate. They formulate a dynamic program to characterize the optimal policy, and build upon the results from Scarf and from Veinott.

This paper differs from the prior work in that we assume an inventory policy, and then characterize its behavior for a class of demand processes. Whereas we have not tried to establish the optimality of this policy, it is a critical-fractile policy, as will be seen. Furthermore, the analysis provides relatively clean results for seeing how the inventory requirements depend upon various problem parameters. We also model a multistage serial system.

2. Single-Item Single-Stage System

In this section we consider a single-item inventory system. The replenishment lead-time is fixed and known, call it L. We assume that the demand process is a nonstationary stochastic process, that the inventory control policy is an adaptive base-stock control policy, and that any demand not satisfied by inventory is backordered. In the following, we describe in more detail the demand process, a forecast model, and an inventory control policy, and then present an analysis of the model. We will introduce additional assumptions as needed.

Demand Process

The demand process is an autoregressive integrated moving average (ARIMA) process given as follows:

$$d_1 = \mu + \varepsilon_1 \quad \text{and}$$
$$d_t = d_{t-1} - (1 - \alpha)\varepsilon_{t-1} + \varepsilon_t$$
$$\text{for } t = 2, 3, \ldots, \tag{1}$$

where d_t is the observed demand in period t, α and μ are known parameters, and $\{\varepsilon_t\}$ is a time series of i.i.d. random variables. We assume that $0 \le \alpha \le 1$, and that ε_t is normally-distributed random noise with $E[\varepsilon_t] = 0$ and $Var[\varepsilon_t] = \sigma^2$. This process is known as an integrated moving average (IMA) process of order (0, 1, 1), and is discussed in detail in Box et al. (1994).

To motivate this demand model, we expand (1) as follows:

$$d_t = \varepsilon_t + \alpha\varepsilon_{t-1} + \alpha\varepsilon_{t-2} + \cdots + \alpha\varepsilon_1 + \mu. \tag{2}$$

Thus we can express the demand as a function of the time series of random noise or independent shocks. Indeed, Muth (1960) provides the following interpretation for (2):

> The shock associated with each time period has a weight of unity; its weight in successive time periods, however, is constant and somewhere between zero and one. Part of the shock in any period therefore, has a permanent effect, while the rest affects the system only in the current period.

As a consequence, each period there is a shift in the mean of the demand process. The shift is proportional to the size of the shock; namely, the shock in period t (ε_t) shifts the mean of the demand process by $\alpha\varepsilon_t$, as of the next period.

By varying α, we can model a range of demand processes. When $\alpha = 0$, the demand follows a stationary i.i.d. process with mean μ, and variance σ^2. For $0 < \alpha \leq 1$, the demand process is a nonstationary process, in which larger values of α result in a less stable or more transitory process; that is, as α grows, d_t depends more and more on the most recent demand realization. We view α as a measure of the inertia in the process; the larger α is, the less inertia there is in the process. When $\alpha = 1$, the demand process is a random walk on a continuous state space; from (1), we see that demand in the next period is the demand in the current period plus a noise term.

This demand process permits negative demand. We note this as a caveat since, in most industrial contexts, negative demand is unlikely or not allowed. Hence, as with any model, some judgment is required as to the applicability of this model of the demand process to the real world. In § 3, we comment on how to assess the likelihood of negative demand.

For an illustrative example of this demand process, we cite Montgomery and Johnson (1976, pp. 217–221). For a demand history for a plastic container, they show how one would identify the demand time series to be an (0,1,1) IMA process and estimate the parameters.

Forecast Model

A first-order exponential-weighted moving average provides the minimum mean square forecast for this demand process (Muth 1960, Box et al. 1994). To see this, consider an exponential-weighted moving average with parameter α and initial forecast μ. We define F_{t+1} to be the forecast, made after observing demand in time period t, for demand in period $t + 1$:

$$F_1 = \mu \quad \text{and}$$

$$F_{t+1} = \alpha d_t + (1 - \alpha)F_t \quad \text{for } t = 1, 2, \ldots. \tag{3}$$

By subtracting equation (3) from (1), one can show by induction that the forecast error is:

$$d_t - F_t = \varepsilon_t \quad \text{for } t = 1, 2, \ldots. \tag{4}$$

Thus, we see that the exponential-weighted moving average is an unbiased forecast and the forecast error is the random noise term for time period t. Hence, there is no better forecast model for this demand process.

Furthermore, consider the class of ARIMA processes that can be expressed as a weighted sum of independent shocks:

$$d_t = \varepsilon_t + \sum_{i=1}^{\infty} w_i \varepsilon_{t-i} + \mu.$$

Muth (1960) shows that the time series given by (1) is the only ARIMA process for which the exponential-weighted moving average provides the best forecast. That is, for all other ARIMA processes, there is a better forecast than (3).

From (4) we see that the variance of the forecast error does not depend upon the magnitude of the demand process. Hence, if we expect the size of the forecast error to be proportional to the demand level, then the (0,1,1) IMA process is probably not a good model for the demand process.

From (3) and (4), we can re-express the forecast in terms of the random noise terms:

$$\begin{aligned} F_{t+1} &= F_t + \alpha\varepsilon_t \\ &= \alpha\varepsilon_t + \alpha\varepsilon_{t-1} + \ldots + \alpha\varepsilon_1 + \mu. \end{aligned} \tag{5}$$

We note that as of time t, the forecast for demand in any future period equals the forecast for the next period, namely F_{t+1}.

Inventory Control Policy

Let q_t be the order placed in period t for delivery in period $t + L$. We assume that in each period t, we first observe d_t, determine this period's order (q_t), receive the order from L periods ago (q_{t-L}), and then fill demand from inventory. Any demand that cannot be met from inventory is backordered. The inventory balance equation for this system is

$$x_t = x_{t-1} - d_t + q_{t-L} \quad \text{for } t = 1, 2, \ldots, \tag{6}$$

where x_t denotes the on-hand inventory (or backorders) at the end of period t. We assume that we can set an initial inventory level x_0, and that $q_t = \mu$ for $t \leq 0$.

Suppose that we operate with a base-stock policy, but adjust the base stock as the demand forecast changes. We propose the following rule to do this:

$$q_t = d_t + L(F_{t+1} - F_t), \tag{7}$$

where F_t is the forecast given by (3). This is effectively the myopic policy (Veinott 1965) for an L-period lead-time, assuming stationary cost parameters. That is, it minimizes the expected one-period cost a lead-time into the future. There are two components to the order quantity. The first component replenishes the demand for the immediate period, as with a typical base-stock policy. The second component adjusts the base-stock level to accommodate the change in the forecast, which changes the mean lead-time demand. In particular, as of time t, the new forecast for demand over the lead-time is LF_{t+1}, rather than LF_t in the prior period $t - 1$.

In posing the ordering policy (7), we permit the order quantity to be negative. Whereas this assumption is unrealistic in most contexts, it does make the analysis of the model much easier. Based on experience with similar modeling assumptions, we expect that the general nature of the results are invariant to this assumption; nevertheless, this should be investigated. Furthermore, in §3, we show that $\{q_t\}$ is an IMA (0,1,1) process with known parameters, and discuss how, for a given value of q_t, one can assess the likelihood that the order quantity becomes negative in the future. In this way, one may judge the applicability of this model of the order process to a given context.

We do not contend that this policy is optimal. Rather, the policy seems reasonable as an extension of the base-stock policy to the case of nonstationary demand. Furthermore, we show below that we can set x_0 to assure that the stock-out probability in each period equals some specified target level. As such, this policy is a critical-fractile policy. Veinott (1965) has shown that if negative orders are permitted, then a critical-fractile policy is optimal for the case of stationary cost parameters. Morton and Pentico (1995) have shown that a critical-fractile policy, while not optimal in general, is near optimal for a class of finite horizon inventory problems with nonstationary demand.

Characterization of Inventory Random Variable

For the given assumptions and control policy, we will demonstrate for $t = 1, 2, \ldots$ that

$$\mathbf{P1} \quad x_t = x_0 - \sum_{i=0}^{L-1} \varepsilon_{t-i} (1 + i\alpha)$$

where $\varepsilon_t = 0$ for $t \le 0$. (We assume the convention $\Sigma_{i=a}^{b} (\ ) = 0$ for $b < a$.)

Proof. First we remove q_{t-L} from (6) by substitution of (7). By repeated backward substitution, we then can rewrite (6) for $t \ge L$ as

$$\begin{aligned} x_t &= x_0 - d_t - d_{t-1} - \cdots - d_{t+1-L} + LF_{t+1-L} \\ &= x_0 - (d_t - F_{t+1-L}) - (d_{t-1} - F_{t+1-L}) \\ &\quad - \cdots - (d_{t+1-L} - F_{t+1-L}). \end{aligned}$$

We use (2) and (5) to express x_t in terms of the random noise terms:

$$\begin{aligned} x_t = x_0 &- (\varepsilon_t + \alpha\varepsilon_{t-1} + \alpha\varepsilon_{t-2} + \cdots + \alpha\varepsilon_{t+1-L}) \\ &- (\varepsilon_{t-1} + \alpha\varepsilon_{t-2} + \alpha\varepsilon_{t-3} + \cdots + \alpha\varepsilon_{t+1-L}) \\ &- \cdots - (\varepsilon_{t+2-L} + \alpha\varepsilon_{t+1-L}) - (\varepsilon_{t+1-L}). \end{aligned}$$

By combining terms, we get the desired result.

For $1 \le t < L$, one can demonstrate the property **P1** in a similar manner by direct substitution and by application of the stated boundary conditions.

Discussion

From **P1** and the assumption that $\{\varepsilon_t\}$ is a time series of normally-distributed i.i.d. random variables with $E[\varepsilon_t] = 0$ and $\mathrm{Var}[\varepsilon_t] = \sigma^2$, we find that x_t is normally distributed with

$$E\,[x_t] = x_0$$

and

$$\begin{aligned} \mathrm{Std}\,[x_t] &= \sigma \sqrt{\sum_{i=0}^{L-1} (1 + i\alpha)^2} \\ &= \sigma \sqrt{L} \sqrt{1 + \alpha(L - 1) + \frac{\alpha^2 (L - 1)(2L - 1)}{6}} \end{aligned} \tag{8}$$

where $t \ge L$ and Std[ ] denotes the standard deviation.

Thus, for the inventory control policy given by (7), the inventory in each period is normally distributed with the same mean and standard deviation. From (8) the standard deviation is a function of the parameters for the demand process and the lead-time; indeed, we can show that $\mathrm{Std}[x_t]$ is the standard deviation of demand over the replenishment lead-time L. The mean, however, is the initial inventory level, which for our purposes is a control variable. We note that x_0 is the safety stock for this inventory system.

In most contexts we would set the safety stock, namely the initial inventory level, to assure some specified level of service. For example, we might set the initial inventory level to be a multiple of the standard deviation of the inventory random variable (which equals the standard deviation of demand over the lead-time):

$$x_0 = z\sigma\sqrt{L}\sqrt{1 + \alpha(L - 1) + \alpha^2 \frac{(L - 1)(2L - 1)}{6}}.$$

In this way, one assures that the probability of not stocking out in each period is $\Phi(z)$, where $\Phi(\ )$ is the cumulative distribution function for the standard normal random variable. Thus, we can set x_0 to achieve any critical-fractile of service.

As one implication of this, suppose that in each time period we incur a per unit inventory holding cost h if $x_t > 0$ or a per unit backorder cost b if $x_t < 0$. Then to minimize the expected costs in each period, one sets x_0 so that the probability of not stocking out equals $b/(h + b)$; that is, in each time period we solve a newsboy problem (e.g., Lee and Nahmias 1993). Thus, the assumed inventory control policy is optimal, provided we permit q_t to be negative (Veinott 1965).

Alternatively, similar to Johnson and Thompson (1975), we could establish conditions on the demand process that assure that q_t is nonnegative. For instance, we can rewrite (7) as

$$q_t = d_t + L\alpha\varepsilon_t.$$

Then we can guarantee that the order quantity is nonnegative if we assume that the random shock is bounded, e.g., $-k\sigma \le \varepsilon_t \le k\sigma$, and that there is a lower bound on demand, equal to $(L\alpha)(k\sigma)$. When these additional assumptions are not appropriate, then I expect that the optimal solution is more complex and would require the solution of a dynamic program. In this case, due to additional constraints on the order quantity, the required safety stock would be greater and the order quantity would be less variable.

The standard deviation of x_t is a surrogate for the safety stock requirements for this inventory system. In the following, we use equation (8) to show how the safety stock depends on the lead-time and on the parameter α, a measure of the inertia of the demand process.

When $\alpha = 0$, we get the familiar result that the safety stock for a stationary, i.i.d. demand process grows with the square root of L (Lee and Nahmias, 1993)

$$\text{Std }[x_t] = \sigma\sqrt{L}.$$

When $\alpha = 1$, the demand process is a random walk and we find

$$\text{Std }[x_t] = \sigma\sqrt{\sum_{i=0}^{L-1} (1 + i)^2} = \sigma\sqrt{\frac{L\ (L + 1)(2L + 1)}{6}}.$$

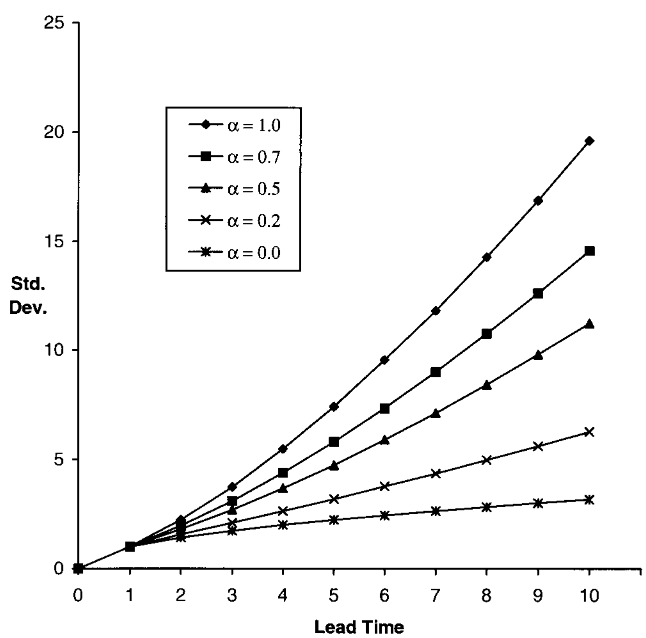
Thus, the safety stock requirement, as a function of the replenishment lead-time, behaves much differently than in the stationary demand case. Instead of having a concave relationship between lead-time and the safety stock, we have a convex relationship; one needs increasing amounts of safety stock as the lead-time grows, when the demand process follows a random walk.

To explore this relationship further, we plot in Figure 1 the standard deviation of x_t as a function of the lead-time for $\sigma = 1$ and various choices of α.

There are two key observations for the assumed demand process from Figure 1.

First, we observe that *we require dramatically more safety stock when demand is nonstationary, in comparison with the textbook case of stationary demand* ($\alpha = 0$). For

Figure 1 **Standard Deviation of x_t as a Function of Lead-Time for $\sigma = 1$**



instance, the safety stock requirements are 50% greater than the stationary case when $L = 6$ for $\alpha = 0.2$, when $L = 3$ for $\alpha = 0.5$, and when $L = 2$ for $\alpha = 1.0$. The safety stock requirements are 100% greater than the stationary case when $L = 10$ for $\alpha = 0.2$, when $L = 5$ for $\alpha = 0.5$, and when $L = 3$ for $\alpha = 1.0$.

Second, we observe that *the relationship between lead-time and safety stock becomes convex for nonstationary demand.* That is, after some point, the rate of increase in safety stock increases with the lead-time. For instance, for $\alpha = 0.2$, the relationship is convex after $L = 3$, for $\alpha = 0.5$ after $L = 2$, and for $\alpha = 1.0$ after $L = 1$. This is in stark contrast to the stationary-demand case where safety stock is a concave function of the lead-time for all values of L.

Erkip et al. (1990) have also observed that significantly more safety stock may be needed when demand does not come from a stationary i.i.d. process. They consider a multi-echelon inventory system in which demand is stationary but is correlated across sites and across time. They develop an explicit expression for the safety stock requirements and show how these requirements are impacted by demand correlation over time.

To this point, the inventory system operates with periodic review where the review period is one time period, however defined. An open question is how does the inventory vary with changes in the length of the review period. In the appendix we extend the model to permit the review period to be a parameter. We develop the analog for equation (8), the standard deviation of the inventory, for the case of continuous review.

3. Single-Item Multi-Stage System: Implications for a Supply Chain

In the previous section we presented an inventory model for a single item and single stage. Suppose now that we have two stages in series, where the prior model and analyses apply to the downstream stage.

Order Amplification

First consider the demand process of the upstream stage, namely the order stream $\{q_t\}$ from the downstream stage. From (2), (5), and (7), we find

$$q_t = (1 + L\alpha)\, \varepsilon_t + \alpha\varepsilon_{t-1} + \alpha\varepsilon_{t-2} + \cdots + \alpha\varepsilon_1 + \mu \quad \text{for } t = 1, 2, \ldots. \tag{9}$$

From (9) we can express the time series $\{q_t\}$ in the standard form for an ARIMA process, namely:

$$q_1 = \mu + \zeta_1 \quad \text{and}$$
$$q_{t+1} = q_t - (1 - \beta)\zeta_t + \zeta_{t+1} \quad \text{for } t = 1, 2, \ldots, \tag{10}$$

where $\zeta_t = (1 + L\alpha)\, \varepsilon_t$ and $\beta = \alpha/(1 + L\alpha)$.

Thus, the demand process seen by the upstream process is also an IMA (0, 1, 1) *process.* From the assumptions for $\{\varepsilon_t\}$, we see that $\{\zeta_t\}$ is i.i.d. normally distributed random noise with $E[\zeta_t] = 0$ and $\text{Var}[\zeta_t] = (1 + L\alpha)^2 \sigma^2$. And $\beta = \alpha/(1 + L\alpha)$ is the parameter for the inertia of the process. Since $\text{Var}[\zeta_t] \geq \text{Var}[\varepsilon_t]$ and $0 \leq \beta \leq \alpha \leq 1$, the upstream demand process is more variable than the downstream demand, but also has more inertia.

The following exponential-weighted moving average provides the minimum mean square forecast for the upstream demand process:

$$G_1 = \mu \quad \text{and}$$
$$G_{t+1} = \beta q_t + (1 - \beta)G_t \quad \text{for } t = 1, 2, \ldots, \tag{11}$$

where G_{t+1} is the forecast for q_{t+1}, the upstream demand in period $t + 1$, made after observing the upstream demand q_t in period t.

From (10) and (11) one can show by induction that

$$q_t = G_t + \zeta_t \quad \text{for } t = 1, 2, \ldots.$$

But by substituting (5) into (9), we also find that

$$q_t = F_t + (1 + L\alpha)\varepsilon_t \quad \text{for } t = 1, 2, \ldots.$$

Thus, since $\zeta_t = (1 + L\alpha)\varepsilon_t$, the downstream forecast is the same as the upstream forecast, namely $G_t = F_t$, and this forecast is an unbiased estimate not only for d_t, but also for q_t.

Furthermore, by comparison with (4), we see that the order stream q_t is more variable than d_t. For a given value of $F_t = G_t$, we have that

$$\text{Std}\,[q_t \mid F_t] = (1 + L\alpha)\sigma = (1 + L\alpha)\ \text{Std}\,[d_t \mid F_t].$$

Thus there is amplification of the exogenous demand process as the downstream stage passes orders to the upstream stage. A simple measure of the amplification is $(1 + L\alpha)$, the ratio of the standard deviation of the downstream order to that of the demand process.

This is an example of the phenomenon termed the "bullwhip effect" in Lee et al. (1997). Indeed, the above result is similar to that presented in Lee et al. for the "demand signal processing" cause for the bullwhip. However, since we assume a different demand process, we find an explicit expression for the increase in order variability.

This phenomenon has also been described by Forrester (1958, 1961) as an example of industrial dynamics. Sterman (1989) explores and documents this phenomenon using the "beer distribution game" as an experimental context.

Hetzel (1993), as part of a supply chain project at the Eastman Kodak Company, also discovered this phenomenon, which he termed the "springboard effect." Hetzel shows how Kodak's safety stock policies could lead to a springboard effect up a supply chain, triggered by a forecast change in customer demand.

Baganha and Cohen (1998) provide empirical evidence for the bullwhip effect, and present a model for explaining the phenomenon as well as for developing mechanisms to mitigate its effects.

Chen et al. (1996) demonstrate for the case of stationary demand how a moving-average forecast can induce a bullwhip effect in a two-stage system, and then quantify the size of the variance amplification. In a related paper, Chen et al. (1997) extend this analysis to an exponential forecast process, and to the case when the demand process has a linear trend. Their findings complement those in the current paper, which assumes a nonstationary demand process.

Cachon (1998) also considers the case of stationary demand but for a two-echelon system consisting of one supplier and multiple retailers. He characterizes the variability of the orders placed upon the supplier and shows how this variability depends on the structure and parameters of the order policy at the retailers.

Order Correlation

The demand process given by (1) is serially correlated. In particular, given a value of F_t, one can show that

$$\mathrm{Cov}[d_{t+j}, d_{t+k} \mid F_t] = (\alpha + j\alpha^2)\sigma^2$$

$$\text{for } 0 \le j < k \text{ and}$$

$$\mathrm{Var}\,[d_{t+j} \mid F_t] = (1 + j\alpha^2)\,\sigma^2 \quad \text{for } 0 \le j. \tag{12}$$

Since $\{q_t\}$ is also an IMA (0, 1, 1) process, it is also serially correlated:

$$\mathrm{Cov}\,[q_{t+j}, q_{t+k} \mid F_t] = (\alpha(1 + L\alpha) + j\alpha^2)\sigma^2$$

$$\text{for } 0 \le j < k \text{ and}$$

$$\mathrm{Var}\,[q_{t+j} \mid F_t] = ((1 + L\alpha)^2 + j\alpha^2)\sigma^2 \quad \text{for } 0 \le j. \tag{13}$$

We observe that the covariances for the time series $\{q_t\}$ are greater than their corresponding terms for $\{d_t\}$, where the difference increases with both the lead-time and the parameter α. This is consistent with the observation that there is more inertia in the upstream demand process than in the downstream demand process.

We use (12) and (13) to assess the likelihood that the demand process and the order process become negative in the future, respectively. For instance, at the end of time period $t - 1$, demand in time period $t + j$, for $j \ge 0$, is a random normal variable with expectation given by the forecast F_t and variance given by (12). Thus, comparing the mean with the standard deviation gives a good proxy for the likelihood that d_{t+j} will be negative.

Upstream Inventory

We now examine the inventory requirements for the upstream stage. Let p_t be the order placed in period t by the upstream stage upon its supplier. The lead-time for replenishment to the upstream stage is K: An order placed in period t is for delivery in period $t + K$. We assume that in each period t, the upstream stage first observes q_t, determines this period's order (p_t), receives the order from K periods ago (p_{t-K}), and then fills the downstream order from inventory. Any demand that cannot be met from inventory is backordered. The inventory balance equation for the upstream stage is

$$y_t = y_{t-1} - q_t + p_{t-K} \quad \text{for } t = 1, 2, \ldots, \tag{14}$$

where y_t denotes the on-hand inventory (or backorders) at the end of period t. We assume an initial inventory level y_0, and that $p_t = \mu$ for $t \le 0$.

Suppose that the upstream stage operates with an adaptive base-stock policy, similar to (7) for the downstream stage:

$$p_t = q_t + K(G_{t+1} - G_t). \tag{15}$$

Again, there are two components to the order quantity. First the upstream stage must replenish the demand for the immediate period, namely the order placed by the downstream stage. Second, the upstream stage adjusts the order quantity to account for the change in the forecast, given by (11). This adjustment is the increase to the base-stock level at the upstream stage due to the change in the forecast. This adjustment increases the amount of inventory on order, namely the pipeline stock, by the change in forecast multiplied by the lead-time.

We assume that the order quantity in (15) either remains nonnegative for any relevant problem instances, or is permitted to be negative.

For the given assumptions and control policy, we now demonstrate for $t = 1, 2, \ldots$, that

P2 $$y_t = y_0 - \sum_{i=0}^{K-1} \varepsilon_{t-i}\,(1 + (L + i)\alpha)$$

where $\varepsilon_t = 0$ for $t \le 0$.

Proof. To show **P2**, we first observe that **P1** applies to the inventory of the upstream stage since the upstream demand process $\{q_t\}$ is an IMA (0, 1, 1) process and since the order policy (15) is structurally the same as (7). Hence from **P1**, we have by direct substitution that

$$\begin{aligned} y_t &= y_0 - \sum_{i=0}^{K-1} \zeta_{t-i}\,(1 + i\beta) \\ &= y_0 - \sum_{i=0}^{K-1} (1 + L\alpha)\varepsilon_{t-i}\left(1 + i\left(\frac{\alpha}{1 + L\alpha}\right)\right) \\ &= y_0 - \sum_{i=0}^{K-1} \varepsilon_{t-i}\,(1 + (L + i)\alpha) \end{aligned}$$

which is the desired result.

Discussion

From **P2** and the assumption that $\{\varepsilon_t\}$ is a time series of normally-distributed i.i.d. random variables with $E[\varepsilon_t] = 0$ and $\mathrm{Var}[\varepsilon_t] = \sigma^2$, we find that for $t \ge K$, y_t is normally distributed with

$$E\,[y_t] = y_0 \quad \text{and}$$

$$\mathrm{Std}\,[y_t] = \sigma\sqrt{\sum_{i=0}^{K-1} (1 + (L + i)\alpha)^2}. \tag{16}$$

For this discussion we assume that the upstream stage intends to provide a high level of service in filling the downstream orders. We note, though, that providing a high level of internal service in such a supply chain may not be a good policy in terms of total inventory costs (e.g., see Graves 1996).

Given the assumption of a high level of internal service, then y_0 is the safety stock for the upstream stage, and the standard deviation of y_t is a surrogate for the upstream safety stock requirements. In the following we discuss several observations from (16) for the assumed nonstationary demand process.

First, since the form of (16) is similar in structure to that for the downstream stage (8), the observations for the downstream stage apply to the upstream stage too. Namely, *we require dramatically more upstream safety stock when demand is nonstationary, in comparison with the textbook case of stationary demand* ($\alpha = 0$); *and the relationship between the upstream lead-time and the upstream safety stock becomes convex for nonstationary demand.*

Second, when $\alpha > 0$, the standard deviation of the upstream inventory (16) depends not only on the upstream lead-time and the inertia of the exogenous demand process, but also on the downstream lead-time. Hence, *for nonstationary demand, the downstream lead-time impacts the safety stock requirements at both the downstream and the upstream stages of a two-stage supply chain.* One implication of this is as a guide for focusing improvement efforts: there may be more impact from reducing the downstream lead-time than the upstream lead-time. Indeed, when $K \ge L$, we can show that a unit reduction in the downstream lead-time will always result in a greater reduction in inventory holding costs than a unit reduction in the upstream lead-time.

As a third observation, for this model *there is not any benefit from providing the upstream stage with additional information about the exogenous demand or about the order process of the downstream stage.* If the downstream stage follows the adaptive base-stock policy, as specified by (7), then the optimal forecast of the upstream demand process is given by (11). The upstream stage needs to know the parameters for this exponential-weighted moving average (or equivalently for the IMA(0, 1, 1) demand process), and needs to observe its demand process $\{q_t\}$. But there is no benefit to the upstream stage from directly observing the demand $\{d_t\}$ at the downstream stage.

For instance, there is no way for the upstream stage to alleviate the bullwhip effect by having more information. Rather, to decrease the amplification of the demand process, there must be a reduction in the downstream lead-time L, or an increase in the inertia in the demand process (smaller α), or a change in the downstream order policy that somehow smoothes the response to a forecast change.

Of course, if the upstream stage does not know the parameters for the customer demand process, then observing the demand $\{d_t\}$ could be of value to estimate these parameters.

We state these three observations for a two-stage serial system. But it should be clear that they extend immediately to an n-stage serial system, as long as each stage has a deterministic lead-time and follows an adaptive base-stock policy. Under these assumptions, for an n-stage serial system, the orders placed by any stage on its supplier will be an IMA (0, 1, 1) process with known parameters. This leads to a fourth observation: *as the order process moves further upstream, it becomes more variable but also has more inertia.*

Inventory Positioning

We conclude this section with an observation on inventory positioning in a serial system. Consider a two-stage supply chain with replenishment lead-times of L and K time periods for the downstream and upstream stages, respectively. Suppose the upstream stage produces an intermediate good, and the downstream stage converts the intermediate good into a finished good.

If we were to only hold safety stock of finished goods, then from (8) the safety stock requirements would be proportional to

$$\sigma \sqrt{\sum_{i=0}^{L+K-1} (1 + i\,\alpha)^2}, \tag{17}$$

since the safety stock must protect against demand variability over the entire lead-time of $L + K$.

But if we were to hold an intermediate safety stock, then the finished goods safety stock would be proportional to

$$\sigma \sqrt{\sum_{i=0}^{L-1} (1 + i\,\alpha)^2}, \tag{18}$$

where we assume for the sake of discussion that the intermediate safety stock will provide a high level of service, so that the lead-time seen by the finished goods stage is L. In effect, the intermediate inventory is a decoupling inventory for the supply chain.

By application of (14), the safety stock of intermediate goods should be proportional to

$$\sigma \sqrt{\sum_{i=0}^{K-1} (1 + (L + i)\alpha)^2} = \sigma \sqrt{\sum_{i=L}^{L+K-1} (1 + i\,\alpha)^2}. \tag{19}$$

We note that the summation within the square root of (17) gets split into (18) and (19) when we decouple the supply chain with an intermediate inventory. If we view the summation as a proxy for the demand variability over the total lead-time, then we see that the downstream (finished goods) inventory is responsible for the first L elements of this summation whereas the upstream (intermediate goods) inventory must cover the remaining K terms.

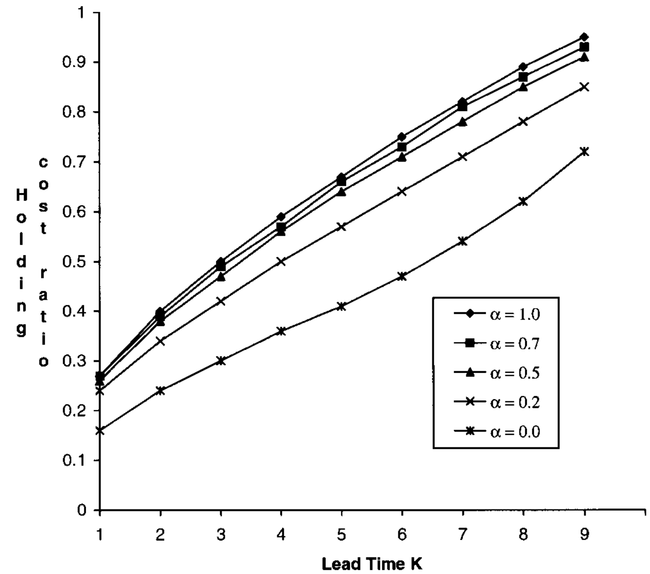
If we were to have more stages and more than one decoupling inventory, this pattern would continue. Given this understanding of how the demand variability gets split up by decoupling inventories, one can easily build a spreadsheet model to explore various strategies for the positioning of decoupling stocks.

As an illustration, we consider a two-stage supply chain for which the ratio of the holding cost for the intermediate good to the holding cost for the finished good is denoted by h. For what value of h are we indifferent between having an intermediate decoupling safety stock versus having no intermediate safety stock? To address this question, we use (17) to approximate the safety-stock holding cost if there is only a finished goods inventory; we use (18) and (19) to approximate the safety-stock holding cost, as a function of h, for the policy with an intermediate decoupling safety stock. In Figure 2 we report the value for h at which we are indifferent between the two stocking options; for this computation, we assume that the total lead-time ($L + K$) is ten, and we plot the indifference value as a function of the intermediate lead-time K for various choices of α.

For instance, if $K = 2$ ($L = 8$) and $\alpha = 0.2$, then we prefer to have an intermediate decoupling safety stock provided the holding cost is less than 34% of the holding cost for the finished good.

Figure 2 provides some insight as to where you would position an intermediate decoupling stock. If one could choose the value of K, then one would only consider locations (values of K) at which the actual holding cost is less than the "breakeven" values in Figure 2. Furthermore, one might suspect that the best choice would be where the difference between the actual holding cost and the breakeven holding cost is greatest. Of course, this conjecture could be examined more closely by using equations (17)–(19).

Figure 2 **The Holding Cost Ratio at Which We Are Indifferent Between Holding a Decoupling Intermediate Safety Stock Versus Holding No Intermediate Safety Stock**



4. Conclusion

In the first part of this paper we present a model for a single-item inventory system with a deterministic lead-time but subject to a stochastic, nonstationary demand process. We propose an adaptive base-stock policy for inventory replenishment and show that it yields a critical-fractile policy. We observe that the safety stock required for the case of nonstationary demand is much greater than for stationary demand; furthermore, the relationship between safety stock and the replenishment lead-time becomes convex when the demand process is nonstationary, quite unlike the case of stationary demand.

The practical significance of this single-item model is that it provides an alternative inventory model for contexts where the assumption of stationary demand is not applicable. In particular this model is suited for items for which an exponentially weighted moving average is an appropriate forecast model. Since many text books on inventory present exponential smoothing as being the model of choice for single-item forecasting, it would seem that the inventory model presented here would have wide applicability (e.g., Brown 1963, Nahmias 1993).

Indeed, in teaching and in consulting, my treatment of forecasting and simple inventory models has often been inconsistent. I have advocated or taught exponential smoothing as a realistic forecast model for many contexts. But whereas this implies that the underlying demand process is nonstationary, I have then assumed a stationary demand process for setting safety stock policies. Possibly this inconsistency is one explanation for why inventory managers often carry more inventory than recommended by textbooks.

The inventory model developed in this paper will at least let me have a consistent story between how we forecast and how we plan safety stocks. Unfortunately, this model will recommend more inventory than what we would typically recommend. A typical practice would be to use the forecast errors from an exponentially smoothed forecast to estimate the variance of the demand process, or equivalently the standard deviation σ; often this has been done by computing the mean absolute deviation (MAD) of the forecast. Given this estimate of σ, we then assume that demand is from a stationary process and recommend safety stock proportional to $\sigma\sqrt{L}$. But if the demand process is truly nonstationary with parameter α, then the safety stock should be larger since the standard deviation of demand over the lead-time is larger, as seen from (8).

In the second part of the paper, we examine how the model extends to a supply chain context. We find that the upstream demand process is also nonstationary and has the same form as for the exogenous downstream demand process. Hence, if the upstream stage also uses an adaptive base-stock policy, then the analysis for the downstream stage applies directly to the upstream-stage inventory. We also observe how the demand process becomes more variable and has more inertia as it is passed back to the upstream stage, and that this bullwhip effect cannot be mitigated by providing more information to the upstream stage.

The two-stage model provides some insight into supply chain behavior. The fact that the analysis of the upstream stage looks the same as for the downstream stage suggests that the single-stage model might serve as a building block for the analysis of more complex systems. But it is somewhat disheartening that if the exogenous demand is nonstationary *as specified in this paper*, then there is no recourse for the upstream stages from the bullwhip effect, given fixed replenishment lead-times. The analysis for the two-stage model does show the importance of lead-time reduction in a supply chain context; in particular we see that reducing the downstream lead-time impacts both the downstream and upstream safety stocks and is likely to provide more benefit than a similar reduction of the upstream lead-time.

There are a lot of unanswered questions or open issues worthy of further research. We examine the simplest of nonstationary processes in the simplest inventory context. It would certainly be of interest to enrich either the model of the demand process or the inventory context or both. For instance, what happens if the replenishment lead-times are stochastic? What happens if forecast errors grow with the magnitude of demand? There is also value from validation of the demand process model in industrial settings, and verification of the applicability of the inventory models. Finally, the optimality or near-optimality of the adaptive base-stock policy is an open question, as well as the implications from allowing replenishment orders to be negative.[1]

Appendix: Asymptotic Behavior of Single-Item Single Stage System

In the paper, the inventory system operates with periodic review where the review period is one time period. In the appendix, we examine how the model extends to reflect changes in the length of the review period.

To investigate this, we develop the model on the time continuum where the length of the review period is $\Delta = 1/n$, for n a positive integer. The analogous model for the demand process for the review period Δ is:

$$d(\Delta) = \mu\Delta + \varepsilon(\Delta) \quad \text{and}$$
$$d(t + \Delta) = d(t) - (1 - \alpha\Delta)\varepsilon(t) + \varepsilon(t + \Delta) \quad \text{for } t = \Delta, 2\Delta, \ldots,$$

where $d(t)$ is the observed demand at time t for $t = \Delta, 2\Delta, \ldots$. Alternatively, one can interpret $d(t)$ to be the demand over the interval $(t - \Delta, t)$. To reflect the change in review period, the initial demand estimate is now $\mu\Delta$, and the inertia parameter is $\alpha\Delta$. The i.i.d. random noise over the interval $(t - \Delta, t)$, given by $\varepsilon(t)$, is normally-distributed with $E[\varepsilon(t)] = 0$ and $\mathrm{Var}[\varepsilon(t)] = \sigma^2 \Delta$.

We can now restate **P1** for the review period Δ to obtain an expression for $x(t)$, the on-hand inventory at time t:

$$x(t) = x(0) - \sum_{i=0}^{nL-1} \varepsilon(t - i\Delta)(1 + i\alpha\Delta),$$

where $\varepsilon(t) = 0$ for $t \le 0$ and $\Delta = 1/n$ for n a positive integer. (Actually the above expression is true as long as nL is a positive integer. When n and nL are not integer, one can define two processes that bound $x(t)$ below and above by running the summation to $\lfloor nL \rfloor - 1$ and to $\lfloor nL \rfloor$ respectively.)

We then find for $t \ge L$ the standard deviation of $x(t)$ to be:

$$\mathrm{Std}\,[x(t)] = \sigma \sqrt{\frac{1}{n}} \sqrt{\sum_{i=0}^{nL-1} \left(1 + i\,\frac{\alpha}{n}\right)^2}$$
$$= \sigma\sqrt{L} \sqrt{1 + \alpha(L - 1/n) + \frac{\alpha^2}{6}(L - 1/n)(2L - 1/n)}. \quad \text{(A1)}$$

Of interest is what happens in the limit as the review period becomes smaller ($n \to \infty$):

$$\mathrm{Std}\,[x(t)] = \sigma\sqrt{L} \sqrt{1 + \alpha L + \frac{\alpha^2 L^2}{3}}. \quad \text{(A2)}$$

The equations (A1) and (A2) show clearly how the safety stock depends upon the replenishment lead-time, the inertia of the demand process (as measured by α), and the length of the review period.

[1]The author wishes to thank David Kletter, Donald Rosenfield, Hau Lee, Steven Nahmias, the senior editor and the referees for their helpful comments on earlier drafts of the paper. The research has been supported in part by the MIT Leaders for Manufacturing Program, a partnership between MIT and major US manufacturing firms.

We note from (A1) that the standard deviation of the inventory increases as we make the review period smaller, suggesting that more safety stock is needed with smaller review periods. This anomalous result is an artifact of how we define the replenishment lead-time and the timing of events. We assume that an order placed at time $t - L$ is available to serve demand at time t. As we partition the time continuum into time intervals of length Δ, we assume that an order placed at the end of the time interval $(t - L - \Delta, t - L)$ is available to serve demand in the time interval $(t - \Delta, t)$. Thus, the order placed at time $t - L$ may serve demand that occurs as early as $t - \Delta$, implying that the lead-time can be as small as $L - \Delta$. Given these assumptions, as we reduce Δ, we actually increase the average lead-time and hence increase the variance of the inventory.

If we were to assume that an order placed at the end of the time interval $(t - L - \Delta, t - L)$ is available to serve demand in the time interval $(t, t + \Delta)$, then we would find (A1) to be:

$$\text{Std}\,[x(t)] = \sigma\sqrt{L + 1/n}\,\sqrt{1 + \alpha L + \frac{\alpha^2 L}{6}(2L + 1/n)}. \qquad \text{(A3)}$$

Equation (A3) has the same limit as (A1), but now the inventory requirements decrease with smaller review periods.

In practice, one's choice of (A1) versus (A3) depends on the context. Nevertheless, the general behavior of the two models is effectively the same. For this presentation, I prefer the model associated with (A1) since it implies that no safety stock is required when $L = 0$.

References

Azoury, K. S. 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Sci.* **31** (9) 1150–1160.

Baganha, M. P., M. A. Cohen. 1998. The stabilizing effect of inventory in supply chains. *Oper. Res.* **46** S72–S83.

Box, G. E. P., G. M. Jenkins, G. C. Reinsel. 1994. *Time Series Analysis Forecasting and Control,* 3rd Ed. Holden-Day, San Francisco, CA. 110–114.

Brown, R. G. 1963. *Smoothing, Forecasting and Prediction of Discrete Time Series.* Prentice Hall, Englewood Cliffs, NJ.

Cachon, G. P. 1998. Managing supply chain demand variability with scheduled ordering policies. Forthcoming *Management Sci.*

Chen, F. Y., J. K. Ryan, D. Simchi-Levi. 1997. The impact of exponential smoothing forecasts on the bullwhip effect. Working Paper, 35 pp.

Chen, F. Y., Z. Dresner, J. K. Ryan, D. Simchi-Levi. 1996. Revised 1998. Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times and information. Working Paper, 41 pp.

Erkip, N., W. H. Hausman, S. Nahmias. 1990. Optimal centralized ordering policies in multi-echelon inventory systems with correlated demands. *Management Sci.* **36**(3) 381–392.

Forrester, J. W. 1958. Industrial dynamics—A major breakthrough for decision making. *Harvard Business Rev.* **36**(4) 37–66.

———. 1961. *Industrial Dynamics.* MIT Press, Cambridge, MA.

Graves, S. C. 1988. Safety stocks in manufacturing systems. *J. Manuf. Oper. Management* **1**(1) 67–101.

———. 1996. A multiechelon inventory model with fixed replenishment intervals. *Management Sci.* **42**(1) 1–18.

Hetzel, W. B. 1993. Cycle time reduction and strategic inventory placement across a multistage process. MIT Master's thesis, Cambridge, MA.

Johnson, G. D., H. E. Thompson. 1975. Optimality of myopic inventory policies for certain dependent demand processes. *Management Sci.* **21**(11) 1303–1307.

Lee, H. L., S. Nahmias. 1993. Single-product, single-location models. S. C. Graves, A. H. Rinnooy Kan and P. H. Zipkin, eds. *Handbooks in Operations Research and Management Science, Vol. 4, Logistics of Production and Inventory.* North-Holland, Amsterdam, Chapter 1.

———, V. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Sci.* **43**(4) 546–558.

Lovejoy, W. S. 1990. Myopic policies for some inventory models with uncertain demand distributions. *Management Sci.* **36**(6) 724–738.

Miller, B. 1986. Scarf's state reduction method, flexibility and a dependent demand inventory model. *Oper. Res.* **34**(1) 83–90.

Montgomery, D. C., L. A. Johnson. 1976. *Forecasting and Time Series Analysis.* McGraw-Hill, New York.

Morton, T. E., D. W. Pentico. 1995. The finite horizon nonstationary stochastic inventory problem: Near-myopic bounds, heuristics, testing. *Management Sci.* **41**(2), 334–343.

Muth, J. F. 1960. Optimal properties of exponentially weighted forecasts. Amer. Statist. Association J. **55** 299–306.

Nahmias, S. 1993. *Production and Operations Analysis.* Irwin, Burr Ridge, IL.

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

———. 1960. Some remarks on Bayes solution to the inventory problem. *Naval Res. Logist.* **7** 591–596.

Song, J., P. Zipkin. 1993. Inventory control in a fluctuating demand environment. *Oper. Res.* **41** 351–370.

Sterman, J. D. 1989. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making framework. *Management Sci.* **35**(3) 321–339.

Veinott, A. 1965. Optimal policy for a multi-product, dynamic, nonstationary inventory problem. *Management Sci.* **12** 206–222.

Accepted by Paul Zipkin; received February 21, 1997. This paper has been with the author for 12 months for 2 revisions.