Analysis Of Ordering Patterns In Serial Supply Chains

John A. Muckstadt

Technical Report 1277

November 2000

1 Introduction

When evaluating the variability of order quantities at each stage of a supply chain compared to true variability in the customer demand on the system, the following questions arise. What impact does the forecasting rule have on the variability of order quantities? How do lead times affect order quantities? How does the use of information affect order quantities? How does the inventory policy affect the order quantity and the variability of the order quantity? The purpose of this note is to examine these questions in certain situations. Specifically, our goal is to demonstrate how the variability of order size is influenced by these factors. We will first examine a two-stage system consisting of a manufacturer and one distribution center (DC). Subsequently, we will consider a supply chain that has several stages through which material must pass before reaching the DC. We will also study two forecasting methods, moving average and exponential smoothing and establish the effect of using these two methods.

2 A Two-Stage system

We begin by examining a two-stage system in which the DC operates as follows. At the beginning of each period, the demand is observed at the DC. If there is sufficient stock on hand, the order is satisfied immediately; otherwise, the demand in excess of supply is

backordered. Next we observe the inventory position at the DC. Let I_t denote this value just prior to the end of period t. Depending on the value of I_t, the length of the lead time, the forecasting mechanism used, and the safety stock requirements at the period's end, the DC places an order on the manufacturing facility for an amount we denote by x_t. Let the lead time from the manufacturing facility to the DC be denoted by τ, which is a known fixed length of time measured in periods. We assume $\tau \geq 1$. We assume that an order placed at the end of period t arrives at the beginning of period $t+\tau$ in time to satisfy the demand in period $t+\tau$. Let us further assume that the demand process in each period at the DC is described by a random variable whose mean is μ and whose variance is σ^2. This distribution remains the same in each period. Furthermore, assume the demands are independent from period to period. Thus the demand process random variables, d_t, are i.i.d.

2.1 Inventory Policies

Let us assume that an order-up-to policy is followed at the DC. That is, if the inventory position is I_t just prior to the end of a period and y_t is the desired order-up-to level at the end of the period, then an amount

$$x_t \;=\; y_t \;-\; I_t \tag{2.1}$$

is ordered. (As we will see, it is possible for $x_t < 0$; however, we will ignore this possibility and its consequences.) But how should y_t be set? There are several possibilities. First, we could set y_t to equal the expected lead time demand plus some number of standard deviations of demand over the lead time. Suppose F_t^τ represents the forecast made in period t after demand occurred in that period for demand over the periods $t+1$ through $t+\tau$. Likewise, let $\hat{S}_t^\tau$ be an estimate of the standard deviation of the forecast error for periods $t+1$ through $t+\tau$.

Thus in this case we would set

$$y_t \;=\; F_t^\tau \;+\; k * \hat{S}_t^\tau, \tag{2.2}$$

where k measures the number of standard deviations of lead time demand that is desired as safety stock. (When demand is normally distributed, this policy is optimal.) Observe that $\hat{S}_t^\tau$ is an estimate of the true forecast error, which we denote by S_t^τ. Suppose σ^τ is the true standard deviation of demand over the lead time τ. Then we can show that $S_t^\tau = constant \cdot \sigma^\tau$, where the value of the constant depends on the forecasting method employed. It is imperative to use S_t^τ rather than σ^τ when calculating the inventory levels. By using S_t^τ we are considering both the effects of demand uncertainty and of forecasting the mean lead time demand.

Another method for setting the order-up-to value, which is commonly used in industry, is as follows:

$$y_t = w \cdot F_t^\tau, \tag{2.3}$$

where $w \geq 1$ reflects the number of periods of forecasted demand that are held in safety stock. We will study the effect of using both policies in this paper.

2.2 Forecasting Demand

There are many ways to estimate both the lead time demand and the standard deviation of the forecast error. Let us first analyze the effect of using one method, the moving average.

Assume the DC uses the demand from the last $m+1$ periods to estimate these statistics. Then

$$F_t^\tau = \tau \cdot \left[\frac{\sum_{j=0}^{m} d_{t-j}}{m+1} \right]. \tag{2.4}$$

Let

$$e_t = d_t - F_{t-1}^1. \tag{2.5}$$

Furthermore, let S_t^2 be the variance of the single period forecast error, that is,

$$S_t^2 \;=\; VAR(e_t) \;=\; VAR(d_t - F_{t-1}^1); \tag{2.6}$$

also

$$(S_t^\tau)^2 \;=\; VAR\left[\sum_{j=1}^{\tau} d_{t+j} \;-\; F_t^\tau\right] \tag{2.7}$$

$$\text{and } S_t^\tau \;=\; \sqrt{VAR\left[\sum_{j=1}^{\tau} d_{t+j} - F_t^\tau\right]}. \tag{2.8}$$

Note that

$$VAR\left\{\sum_{j=0}^{m} d_{t-j}\right\} \;=\; (m+1)\cdot\sigma^2 \tag{2.9}$$

and that

$$\begin{aligned}
VAR\left\{d_t - F_{t-1}^1\right\} &= VAR(d_t) \;+\; VAR(F_{t-1}^1) && (2.10)\\
&= \sigma^2 \;+\; VAR\left\{\frac{\sum_{j=1}^{m+1} d_{t-j}}{m+1}\right\} && (2.11)\\
&= \sigma^2 \;+\; \frac{1}{(m+1)^2}\cdot(m+1)\cdot\sigma^2 && (2.12)\\
&= \sigma^2\left(1+\frac{1}{m+1}\right) \;=\; \sigma^2\frac{(m+2)}{(m+1)}. && (2.13)
\end{aligned}$$

The variance of the forecast error is

$$\begin{aligned}
(S_t^\tau)^2 \;=\; VAR\left(\sum_{j=1}^{\tau} d_{t+j} - F_t^\tau\right) &= VAR\left(\sum_{j=1}^{\tau} d_{t+j} - \frac{\tau}{m+1}\cdot\sum_{j=0}^{m} d_{t-j}\right) && (2.14)\\
&= VAR(\sum_{j=1}^{\tau} d_{t+j}) \;+\; \frac{\tau^2}{(m+1)^2}VAR(\sum_{j=0}^{m} d_{t-j}) && (2.15)\\
&= \tau\sigma^2 \;+\; \frac{\tau^2}{m+1}\sigma^2 && (2.16)\\
&= \left(\tau+\frac{\tau^2}{m+1}\right)\sigma^2 = \left(1+\frac{\tau}{m+1}\right)\tau\sigma^2. && (2.17)
\end{aligned}$$

Observe that, depending on the magnitude of the term $\frac{\tau}{(m+1)}$, the variance of the forecast error over the lead time could be substantially greater than the variance of the demand

process over τ periods, $\sigma^2\tau$. Consequently, it is essential to use the standard deviation of the forecast errors as the basis for determining safety stock levels.

The variance and standard deviation of the forecast error over the lead time must, of course, be estimated. Using the demand observations over the $m+1$ periods, we construct this estimate as

$$\hat{S}_t^\tau = \sqrt{\left(1+\frac{\tau}{m+1}\right)\tau}\sqrt{\frac{\sum_{j=0}^{m}(e_{t-j})^2}{m}}. \tag{2.18}$$

2.3 Measuring the Variance Of The Order Quantity

The variance of the DC order quantity differs from the variance of the demand process. It is always at least as large as σ^2. The question is how much larger can it be. This difference between the two variances is sometimes called the bullwhip effect. Suppose the order-up-to target was at some value y and never changed. Then x_t in each period t would equal the demand in period t, d_t. However, since the forecast of the expected demand and standard deviation of demand over the lead time both change from period-to-period, the order-up-to level is also adjusted each period. Therefore, the order quantity in period t must reflect two things, the demand in that period and the change in the target order-up-to levels. That is,

$$x_t = (y_t - y_{t-1}) + d_t. \tag{2.19}$$

Note that if y_{t-1} is larger than y_t then it is possible for x_t to be negative. Let us assume that in this case that the excess is returned to the manufacturer. Suppose the order-up-to target is calculated as

$$y_t = F_t^\tau + k \cdot \hat{S}_t^\tau. \tag{2.20}$$

Then

$$x_t = (F_t^\tau + k\hat{S}_t^\tau) - (F_{t-1}^\tau + k\hat{S}_{t-1}^\tau) + d_t \tag{2.21}$$

$$= F_t^\tau - F_{t-1}^\tau + k(\hat{S}_t^\tau - \hat{S}_{T_1}^\tau) + d_t \tag{2.22}$$

$$= \frac{\tau \sum_{j=0}^{m} d_{t-j}}{m+1} - \frac{\tau \sum_{j=1}^{m+1} d_{t-j}}{m+1} + k(\hat{S}_t^{\tau} - S_{t-1}^{\hat{\tau}}) + d_t \tag{2.23}$$

$$= \frac{\tau[d_t - d_{t-(m+1)}]}{m+1} + d_t + k(\hat{S}_t^{\tau} - S_{t-1}^{\hat{\tau}}) \tag{2.24}$$

$$= \left(1 + \frac{\tau}{m+1}\right) d_t - \frac{\tau}{m+1} \cdot d_{t-(m+1)} + k\left(\hat{S}_t^{\tau} - S_{t-1}^{\hat{\tau}}\right) \tag{2.25}$$

and

$$VAR(x_t) = \left(1 + \frac{\tau}{m+1}\right)^2 \sigma^2 + \left(\frac{\tau}{m+1}\right)^2 \sigma^2 + k^2 VAR(\hat{S}_t^{\tau} - S_{t-1}^{\hat{\tau}}) \tag{2.26}$$

$$+ 2k\left(1 + \frac{\tau}{m+1}\right) COV(d_t, \hat{S}_t^{\tau}) + 2k \cdot \frac{\tau}{m+1} \cdot COV(d_{t-(m+1)}, S_{t-1}^{\hat{\tau}}) \tag{2.27}$$

$$= \left(1 + 2\frac{\tau}{m+1} + \frac{2\tau^2}{(m+1)^2}\right) \sigma^2 + k^2 VAR(\hat{S}_t^{\tau} - S_{t-1}^{\hat{\tau}}) \tag{2.28}$$

$$+ 2k\left(1 + \frac{2\tau}{m+1}\right) COV(d_t, \hat{S}_t^{\tau}) , \tag{2.29}$$

because the demand process is i.i.d.

Let us make one more assumption. Assume that the probability distribution for the demand process is symmetric about μ, the expected demand per period. Then $COV(d_t, \hat{S}_t^{\tau}) = 0$, where $\hat{S}_t^{\tau} = \sqrt{\left(1 + \frac{\tau}{m+1}\right)\tau}\sqrt{\frac{\sum_{j=0}^{m}(e_{t-j})^2}{m}}$. Let us see why this is the case. Recall that $e_t = d_t - F_{t-1}^1$.

Since d_t is assumed to be symmetric, we can write this random variable as

$$d_t = \mu + \epsilon_t, \tag{2.30}$$

where

$$E(\epsilon_t) = 0 \quad \text{and} \quad VAR(\epsilon_t) = \sigma^2 . \tag{2.31}$$

Then

$$e_t = \mu + \epsilon_t - F_{t-1}^1 = \mu + \epsilon_t - \frac{\sum_{i=1}^{m+1} d_{t-i}}{m+1} \tag{2.32}$$

$$= \mu + \epsilon_t - \sum_{i=1}^{m+1} \frac{(\mu + \epsilon_{t-i})}{m+1} \tag{2.33}$$

$$= \epsilon_t - \frac{1}{m+1} \sum_{i=1}^{m+1} \epsilon_{t-i} \tag{2.34}$$

$$= \sum_{i=0}^{m+1} a_{t-i}\epsilon_{t-i} \tag{2.35}$$

where

$$a_t = 1 \text{ and } a_{t-i} = -\frac{1}{m+1} \text{ for } i = 1, \ldots, m+1. \tag{2.36}$$

Consequently,

$$\hat{S}_t^\tau = \sqrt{\left(1 + \frac{\tau}{m+1}\right)\tau} \cdot \sqrt{\frac{\sum_{j=0}^{m}(e_{t-j})^2}{m}} \tag{2.37}$$

$$= \sqrt{\left(1 + \frac{\tau}{m+1}\right)\tau} \cdot \sqrt{\frac{\sum_{j=0}^{m}\left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i-j}\right)^2}{m}} \tag{2.38}$$

and

$$COV(d_t, \hat{S}_t^\tau) = COV(\epsilon_t, \sqrt{\left(1 + \frac{\tau}{m+1}\right)\tau} \cdot \sqrt{\frac{\sum_{j=0}^{m}\left(-\sum_{i=1}^{m+1}\frac{1}{m+1}\epsilon_{t-i-j} + \epsilon_{t-j}\right)^2}{m}}) \tag{2.39}$$

$$= \sqrt{\left(1 + \frac{\tau}{m+1}\right)\tau} \cdot E\left[\epsilon_t \sqrt{\frac{\sum_{j=0}^{m}\left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i-j}\right)^2}{m}}\right] \tag{2.40}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \epsilon_t \sqrt{\frac{\sum_{j=0}^{m}\left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i-j}\right)^2}{m}} f(\epsilon_t) f(\epsilon_{t-1}) \ldots d\epsilon_1 \ldots d\epsilon_t \ , \tag{2.41}$$

where $f(\epsilon_t)$ measures the density function for the random variable ϵ_t. Recall that $f(\epsilon_t) = f(-\epsilon_t)$. Hence,

$$\begin{aligned}
G &= \int_{-\infty}^{\infty} \epsilon_t \sqrt{\frac{\sum_{j=0}^{m}\left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i-j}\right)^2}{m}} f(\epsilon_t)\, d\epsilon_t \\
&= \int_{0}^{\infty} \epsilon_t \sqrt{\frac{\sum_{j=0}^{m}\left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i-j}\right)^2}{m}} f(\epsilon_t)\, d\epsilon_t \ + \\
&\quad \int_{-\infty}^{0} \epsilon_t \sqrt{\frac{\sum_{j=0}^{m}\left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i-j}\right)^2}{m}} f(\epsilon_t)\, d\epsilon_t \\
&= \int_{0}^{\infty} \epsilon_t \sqrt{\frac{\sum_{j=0}^{m}\left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i-j}\right)^2}{m}} f(\epsilon_t)\, d\epsilon_t \\
&\quad - \int_{0}^{\infty} \epsilon_t \sqrt{\frac{\left(-\epsilon_t - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i}\right)^2 + \sum_{j=1}^{m}\left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1}\epsilon_{t-i-j}\right)^2}{m}} f(\epsilon_t)\, d\epsilon_t
\end{aligned}$$

Multiplying and dividing by the sum of the two square root terms yields

$$G = \int_0^\infty \frac{\epsilon_t \left\{ \sum_{j=0}^m \frac{(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1} \epsilon_{t-i-j})^2}{m} - \frac{(-\epsilon_t - \frac{1}{m+1}\sum_{i=1}^{m+1} \epsilon_{t-i})^2}{m} - \frac{\sum_{j=1}^m (\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1} \epsilon_{t-i-j})^2}{m} \right\}}{\sqrt{\frac{\sum_{j=0}^m (\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1} \epsilon_{t-j})^2}{m}} + \sqrt{\frac{(-\epsilon_t - \frac{1}{m+1}\sum_{i=1}^{m+1} \epsilon_{t-i})^2 + \sum_{j=1}^m (-\epsilon_{t-j} - \frac{\sum_{i=1}^{m+1} \epsilon_{t-j}}{m+1})^2}{m}}} f(\epsilon_t)\, d\epsilon_t$$

But

$$\sum_{j=0}^m \left(\epsilon_{t-j} - \frac{1}{m+1} \sum_{i=1}^{m+1} \epsilon_{t-i-j} \right)^2 - \left(\epsilon_t + \frac{1}{m+1} \sum_{i=1}^{m+1} \epsilon_{t-i} \right)^2 - \sum_{j=1}^m \left(\epsilon_{t-j} - \frac{1}{m+1} \sum_{i=1}^{m+1} \epsilon_{t-i-j} \right)^2 =$$

$$\epsilon_t^2 - \frac{2\epsilon_t}{m+1} \sum_{i=1}^{m+1} \epsilon_{t-i} + \left(\frac{1}{m+1} \right)^2 \left(\sum_{i=1}^{m+1} \epsilon_{t-i} \right)^2 - \epsilon_t^2 - \frac{2\epsilon_t}{m+1} \sum_{i=1}^{m+1} \epsilon_{t-i} - \left(\frac{1}{m+1} \right)^2 \left(\sum_{i=1}^{m+1} \epsilon_{t-i} \right)^2 =$$

$-\frac{4\epsilon_t}{m+1} \sum_{i=1}^{m+1} \epsilon_{t-i}$ so that

$$G = -\frac{4}{m+1} \int_0^\infty \frac{\epsilon_t^2 \cdot \sum_{i=1}^{m+1} \epsilon_{t-i} \qquad f(\epsilon_t)\, d\epsilon_t}{\sqrt{\frac{\sum_{j=0}^m \left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1} \epsilon_{t-i-j}\right)^2}{m}} - \sqrt{\frac{\left(\epsilon_t + \frac{1}{m+1}\sum_{i=1}^{m+1} \epsilon_{t-i}\right)^2 + \sum_{j=1}^m \left(\epsilon_{t-j} - \frac{1}{m+1}\sum_{i=1}^{m+1} \epsilon_{t-i-j}\right)^2}{m}}}$$

Consider G as a function of $\epsilon_{t-1}, \ldots, \epsilon_{t-m-1}$. Observe that

$$G(\epsilon_{t-1}, \ldots, \epsilon_{t-m-1}) \quad = \quad -G(-\epsilon_{t-1}, \ldots, -\epsilon_{t-m-1}). \tag{2.42}$$

Consequently, $COV(\epsilon_t, \hat{S}_t^\tau) = 0$ and $COV(d_t, \hat{S}_t^\tau) = 0$, too.

Thus when d_t has a symmetric distribution around μ we see that

$$VAR(x_t) \quad = \quad \left(1 + 2\frac{\tau}{m+1} + 2\frac{\tau^2}{(m+1)^2} \right) \sigma^2 + k^2 VAR(\hat{S}_t^\tau - \hat{S}_{t-1}^\tau). \tag{2.43}$$

Since $k^2 VAR(\hat{S}_t^\tau - \hat{S}_{t-1}^\tau) \geq 0$, we have the following.

Theorem

$$\frac{VAR(x_t)}{\sigma^2} \quad \geq \quad 1 + \frac{2\tau}{m+1} + \frac{2\tau^2}{(m+1)^2} \tag{2.44}$$

whenever the demand distribution is i.i.d and symmetric at the DC.

Clearly the bound depends on both τ and m. For τ large relative to m we see that the

variance of the order quantity can be substantially larger than the variance of the demand process. The manufacturer's order process is therefore quite sensitive to both τ and m. By keeping the lead time low and the ratio of $\frac{\tau}{m+1}$ low, the variation in the order quantity will more closely match the demand process variability.

Suppose we additionally assume that demands are normally distributed in all periods. Then, after much analysis, we can prove the following.

Theorem

$$\frac{VAR(x_t)}{\sigma^2} \geq 1 + \frac{2\tau}{m+1} + \frac{2\tau^2}{(m+1)^2} + \frac{k^2\tau(1+\frac{\tau}{m+1})}{(m+1)^2} \tag{2.45}$$

when demands at the DC are i.i.d and normally distributed and the supply chain operates using a moving average forecaster and in the dynamic manner we have stipulated.

Thus we see that if k is large, the amplification of the variance of the order quantity can b quite substantial if $\frac{\tau}{m+1}$ is large. When $\frac{\tau}{m+1}$ is small and $m+1$ is large, the effect of safety stock is minimal. For example, if $\tau = 2$, $m+1 = 10$, and $k^2 = 2.5$, then the lower bound on $\frac{VAR(x_t)}{\sigma^2}$ is 1.275 and the value of the safety stock term is only 0.055 . If $\tau = m+1 = 5$, then the value of the safety stock term is 1 and the lower bound is 6. For fixed $\frac{\tau}{m+1}$, as $m+1$ decreases, the effect of the safety stock term increases.

3 Measuring The Value Of Information

The popular business literature has recently emphasized the importance of sharing information throughout a supply chain. The stated goal is to provide customer demand information to each echelon of the supply chain in each period. The feeling is that by sharing this information, the bullwhip effect will be mitigated. The purpose of this section is to establish what the impact of sharing demand information is and to see if the bullwhip effect can be eliminated. We will examine two cases. First, we study the supply chain system when in-

formation concerning demand is shared; second, customer demand information is not shared among supply chain partners.

3.1 Measuring The Variation In Order Sizes When Customer Demand Information Is Shared Throughout An N Echelon Supply Chain

Suppose we have a supply chain consisting of N stages or echelons. Furthermore, assume there is but one installation at each echelon, that is the system is a serial system. To begin our analysis we assume that demand is observed each period at the DC and is immediately shared throughout the other $N-1$ installations. Furthermore, each echelon in this system uses an $m+1$ period moving average to estimate the expected demand per period. Thus in period t, the forecast of future demand made at each location is $F_t^1 = \frac{\sum_{j=0}^{m} d_{t-j}}{m+1}$.

We assume the supply chain operates as follows. At the beginning of period t, inventory is received at each installation. Customer orders are received for a total of d_t units at the DC and are satisfied if possible; otherwise, the unsatisfied demands are backordered. The DC then recalculates its order-up-to level for period t, which we denote by y_t^1, and orders an amount of x_t^1 units that is required to raise its inventory level to y_t^1. Echelon 2 receives the order for x_t^1 units. Knowing x_t^1 and d_t, echelon 2 calculates its order-up-to level, y_t^2, and places an order for x_t^2 units. This process continues throughout all N stages.

To this point we have not specified what ordering policy is followed at each location at each period. We will consider two policies. In the first policy, we assume that each echelon follows an order-up-to policy based on its installation view of inventory. Specifically, we assume that for each echelon l,

$$y_t^l = \tau_l F_t^1. \tag{3.46}$$

In the second policy, we assume that

$$y_t^l = \left\{ \sum_{u=1}^{l} \tau_u \right\} F_t^1 . \tag{3.47}$$

Thus in this case the policy is based on echelon stock. Note that in both cases we assume the safety stock is zero throughout the system, that is, $k = 0$. We will return to this point subsequently.

3.2 Case 1: Ordering Policy $y_t^l = \tau_l F_t^1$

As before, assume d_t is i.i.d with mean μ and variance σ^2 per period at the DC. Then

$$\begin{aligned} x_t^1 &= y_t^1 - y_{t-1}^1 + d_t & (3.48) \\ &= \left(1 + \frac{\tau_1}{m+1}\right) d_t - \frac{\tau_1}{m+1} \cdot d_{t-(m+1)} , & (3.49) \end{aligned}$$

as we showed earlier. We also showed that

$$VAR(x_t^1) = \left[1 + \frac{2\tau_1}{m+1} + 2\left(\frac{\tau_1}{m+1}\right)^2 \right] \sigma^2 \tag{3.50}$$

when $k = 0$. By assumption,

$$\begin{aligned} x_t^l &= y_t^l - y_{t-1}^l + x_t^{l-1} & (3.51) \\ &= y_t^l - y_{t-1}^l + y_t^{l-1} - y_{t-1}^{l-1} + x_t^{l-2} & (3.52) \\ &= y_t^l - y_{t-1}^l + y_t^{l-1} - y_{t-1}^{l-1} + \ldots + y_t^1 - y_{t-1}^1 + d_t & (3.53) \\ &= \tau_l \cdot (F_t^1 - F_{t-1}^1) + \tau_{l-1}(F_t^1 - F_{t-1}^1) + \ldots + \tau_1(F_t^1 - F_{t-1}^1) + d_t & (3.54) \\ &= \left\{ \sum_{u=1}^{l} \tau_u \right\} (F_t^1 - F_{t-1}^1) + d_t & (3.55) \\ &= \left(1 + \frac{\sum_{u=1}^{l} \tau_u}{m+1} \right) d_t - \left\{ \frac{\sum_{u=1}^{l} \tau_u}{m+1} \right\} d_{t-(m+1)} , \ l \geq 2 . & (3.56) \end{aligned}$$

In section 2.3, we showed that (2.43) holds. Replacing the constant τ with $\sum_{u=0}^{l} \tau_u$ we see that

$$\frac{VAR(x_t^l)}{\sigma^2} = \left[1 + 2\left(\frac{\sum_{u=1}^{l} \tau_u}{m+1} \right) + 2\left(\frac{\sum_{u=1}^{l} \tau_u}{m+1} \right)^2 \right] . \tag{3.57}$$

We note that the right-hand side provides a lower bound estimate of the left-hand side of the equation when k is greater than zero as shown in equation (2.44).

3.3 Case 2: Echelon Stock Based Ordering Policy

Now let us assume that each location calculates its order-up-to level based on echelon stock. That is

$$y_t^l = \left\{ \sum_{u=1}^{l} \tau_u \right\} \cdot F_t^1 \tag{3.58}$$

In this case, the order quantity at stage l satisfies

$$\begin{aligned} x_t^l &= y_t^l - y_{t-1}^l + d_t \ . && (3.59) \\ &= \left\{ \sum_{u=1}^{l} \tau_u \right\} F_t^1 - \left\{ \sum_{u=1}^{l} \tau_u \right\} F_{t-1}^1 + d_t && (3.60) \\ &= \left\{ 1 + \frac{\sum_{u=1}^{l} \tau_u}{m+1} \right\} d_t - \left\{ \frac{\sum_{u=1}^{l} \tau_u}{m+1} \right\} d_{t-(m+1)} \ . && (3.61) \end{aligned}$$

However, this is identical to 3.56. Hence 3.57 holds in this case as well.

Let us make a couple of important observations. First, $VAR(x_t) > \sigma^2$ so that the bullwhip effect will not be eliminated by communicating demand information throughout the supply chain. Thus even when the simplest ordering policy is followed and the demands are estimated using the same method at each stage of the supply chain, the variance of the order quantity exceeds the variance of the demand process. As the value of $\sum_{u=1}^{l} \tau_u$ increases for fixed m, the $VAR(x_t^l)$ increases. In fact, if we let $T = \sum_{u=1}^{l} \tau_u$, then

$$\frac{d(VAR(x_t^l))}{dT} = \left[\frac{2}{m+1} + \frac{4T}{(m+1)^2} \right] \sigma^2 \ . \tag{3.62}$$

Earlier, we observed that in practice another policy is often followed to create safety stock. In this case, the order-up-to level is set as follows:

$$y_t^l = w_l \cdot \tau_l \cdot F_t^1 \ , \ w_l \geq 1 \tag{3.63}$$

at echelon l. The factor w_l is set to reflect the number of periods of safety stock desired for echelon l. We note that as w_l increases for each echelon l, then the variance of x_t^l increases too. We also observe that echelon l can not set its order-up-to level to

$$y_t^l = w_l \cdot \sum_{u=1}^{l} \tau_u \cdot F_t^1 \tag{3.64}$$

It must be calculated as

$$y_t^l = \{\sum_{u=1}^{l} w_u \cdot \tau_u\} \cdot F_t^1. \tag{3.65}$$

Otherwise the pipeline will not contain enough inventory to support the desired service level.

3.4 Measuring Order Size Variation When No Demand Information Is Shared In The Supply Chain

Suppose an information system structure exists in which only orders are passed from one echelon to the next. Demand information is known only at the DC. Our objective is to ascertain what impact this type of information system has on the variance of the order size at each echelon.

Let us assume that echelon l sets its own policy as follows:

$$y_t^l = \tau_l \cdot F_{t,l} \tag{3.66}$$

$$\text{where } F_{t,1} = \frac{\sum_{j=0}^{m} d_{t-j}}{m+1}, \tag{3.67}$$

$$\text{and } F_{t,l} = \frac{\sum_{j=0}^{m} x_{t-j}^{l-1}}{m+1}. \tag{3.68}$$

As we know,

$$VAR(x_t^1) = \left(1 + \frac{2\tau_1}{m+1} + 2\left(\frac{\tau_1}{m+1}\right)^2\right)\sigma^2 \tag{3.69}$$

$$= \left[\left(1 + \frac{\tau_1}{m+1}\right)^2 + \left(\frac{\tau_1}{m+1}\right)^2\right]\sigma^2. \tag{3.70}$$

Suppose $l = 2$. Then

$$x_t^2 = y_t^2 - y_{t-1}^2 + x_t^1 \tag{3.71}$$

$$= \tau_2 \frac{\sum_{j=0}^{m} x_{t-j}^1}{m+1} - \tau_2 \frac{\sum_{j=1}^{m+1} x_{t-j}}{m+1} + x_t^1 \tag{3.72}$$

$$= \left(1 + \frac{\tau_2}{m+1}\right) x_t^1 - \frac{\tau_2}{m+1} \cdot x_{t-(m+1)}^1 \tag{3.73}$$

But

$$x_t^1 = \left(1 + \frac{\tau_1}{m+1}\right) d_t - \frac{\tau_1}{m+1} \cdot d_{t-(m+1)} \tag{3.74}$$

$$\text{and } x_{t-(m+1)}^1 = \left(1 + \frac{\tau_1}{m+1}\right) d_{t-(m+1)} - \frac{\tau_1}{m+1} d_{t-2(m+1)} \tag{3.75}$$

Let $M_1 = \frac{\tau_1}{m+1}$ and $M_2 = \frac{\tau_2}{m+1}$. Then

$$x_t^2 = (1+M_2)\left[(1+M_1)d_t - M_1 d_{t-(m+1)}\right] - M_2\left[(1+M_1)d_{t-(m+1)} - M_1 d_{t-2(m+1)}\right] \tag{3.76}$$

$$= (1+M_1)(1+M_2)d_t - M_1(1+M_2)d_{t-(m+1)} - (1+M_1)M_2 d_{t-(m+1)} + M_1 M_2 d_{t-2(m+1)} \tag{3.77}$$

$$= (1+M_1)(1+M_2)d_t - [M_1(1+M_2) + (1+M_1)M_2]d_{t-(m+1)} + M_1 M_2 d_{t-2(m+1)} \tag{3.78}$$

$$= (1 + (M_1+M_2) + M_1M_2)d_t - (M_1 + M_2 + 2M_1M_2)d_{t-(m+1)} + M_1M_2 d_{t-2(m+1)}. \tag{3.79}$$

Hence

$$VAR(x_t^2) = (1 + (M_1+M_2) + M_1M_2)^2\sigma^2 + ((M_1+M_2) + 2M_1M_2)^2\sigma^2 + (M_1M_2)^2\sigma^2 \tag{3.80}$$

$$= \left[1 + 2(M_1+M_2) + 2M_1M_2 + (M_1+M_2)^2 + (M_1M_2)^2 + 2(M_1+M_2)M_1M_2\right]\sigma^2 \tag{3.81}$$

$$+ \left[(M_1+M_2)^2 + 4M_1M_2(M_1+M_2) + 4(M_1M_2)^2 + (M_1M_2)^2\right]\sigma^2 \tag{3.82}$$

$$= \left[1 + 2(M_1+M_2) + 2M_1M_2 + 2(M_1+M_2)^2 + 6(M_1M_2)^2 + 6M_1M_2(M_1+M_2)\right]\sigma^2$$

$$= \left[1 + 2(M_1+M_2) + 2(M_1^2+M_2^2) + 6M_1M_2 + 6(M_1M_2)^2 + 6M_1M_2(M_1+M_2)\right]\sigma^2$$

$$> \left[1 + 2(M_1+M_2) + 2(M_1^2+M_2^2) + 4M_1M_2 + 4(M_1M_2)^2 + 4M_1M_2(M_1+M_2)\right]\sigma^2$$

$$= \left[1 + 2M_1 + 2M_1^2\right]\left[1 + 2M_2 + 2M_2^2\right]\sigma^2 \tag{3.83}$$

$$= \prod_{u=1}^{2}\left[1 + 2M_u + 2M_u^2\right]\sigma^2. \tag{3.84}$$

By induction, assume

$$VAR(x_t^l) > \prod_{u=1}^{l}\left[1 + 2M_u + 2M_u^2\right]\sigma^2. \tag{3.85}$$

Then we can show that

$$VAR(x_t^{l+1}) \quad > \quad \prod_{u=1}^{l+1} \left[1 + 2M_u + 2M_u^2\right] \sigma^2, \text{ too.} \tag{3.86}$$

As for the case $l = 2$,

$$x_t^{l+1} \quad = \quad (1 + M_{l+1})x_t^l - M_{l+1} \cdot x_{t-(m+1)}^l \tag{3.87}$$

where $M_l = \frac{\tau_l}{m+1}$ and $M_{l+1} = \frac{\tau_{l+1}}{m+1}$. Then

$$\begin{aligned}
VAR(x_t^{l+1}) &= (1 + M_{l+1})^2 VAR(x_t^l) + (M_{l+1})^2 VAR(x_{t-(m+1)}^l) && (3.88)\\
&\quad -2(1 + M_{l+1})M_{l+1} COV(x_t^l, x_{t-(m+1)}^l) && (3.89)\\
&\geq \left[(1 + M_{l+1})^2 + (M_{l+1})^2\right] VAR(x_{t-(m+1)}) && (3.90)\\
&> \left[(1 + M_{l+1})^2 + (M_{l+1})^2\right] \prod_{u=1}^{l} \left[1 + 2M_u + 2M_u^2\right] \sigma^2 && (3.91)\\
&= \prod_{u=1}^{l+1} \left[1 + 2M_u + 2M_u^2\right] \sigma^2. && (3.92)
\end{aligned}$$

since $COV(x_t^l, x_{t-(m+1)}^l) < 0$, as we shall now see.

To begin,

$$\begin{aligned}
COV(x_t^1, x_{t-(m+1)}^1) &= COV((1 + M_1)d_t - M_1 d_{t-(m+1)}, (1 + M_1)d_{t-(m+1)} - M_1 d_{t-2(m+1)})\\
&= -(1 + M_1)M_1 \cdot VAR(d_{t-(m+1)}) \ < \ 0.
\end{aligned}$$

For $l = 2$,

$$\begin{aligned}
COV(x_t^2, x_{t-(m+1)}^2) &= COV((1 + M_2)x_t^1 - M_2 x_{t-(m+1)}^1, (1 + M_2)x_{t-(m+1)}^1 - M_2 x_{t-2(m+1)}^1)\\
&= (1 + M_2)^2 COV(x_t^1, x_{t-(m+1)}^1) + (M_2)^2 COV(x_{t-(m+1)}^1, x_{t-2(m+1)}^1)\\
&\quad -(1 + M_2)M_2 \cdot VAR(x_{t-(m+1)}^1)\\
&= \left[(1 + M_2)^2 + M_2^2\right] COV(x_t^1, x_{t-(m+1)})\\
&\quad -(1 + M_2)M_2 \cdot VAR(x_{t-(m+1)}^1) \ < \ 0
\end{aligned}$$

since the demand process is stationary.

Next we show that $COV(x_t^l, x_{t-(m+1)}^l) < 0$, $l \geq 2$. Since we know that $COV(x_t^2, x_{t-(m+1)}^2) <$

0, we assume the result through echelon $l-1$ and show the desired result for l.

Clearly,

$$\begin{aligned} COV(x_t^l, x_{t-(m+1)}^l) &= COV((1+M_l)x_t^{l-1} - M_l x_{t-(m+1)}, (1+M_l)x_{t-(m+1)}^{l-1} - M_l x_{t-2(m+1)}^{l-1}) \\ &= \left[(1+M_l)^2 + M_l^2\right] COV(x_t^{l-1}, x_{t-(m+1)}^{l-1}) - (1+M_l)M_l VAR(x_{t-(m+1)}^l) \\ &< 0. \end{aligned} \tag{3.93}$$

Thus we have proven the desired result.

3.5 Evaluating The Effect Of Information Sharing

In the preceding sections we developed lower bounds on the ratio $VAR(x_t^l)/\sigma^2$ for the two cases where demand information is and is not shared throughout the supply chain. When demand information is supplied throughout the N echelons, the ratio is given by

$$\frac{VAR(x_t^l)}{\sigma^2} = 1 + \frac{2\sum_{u=1}^l \tau_u}{m+1} + 2\left(\frac{\sum_{u=1}^l \tau_u}{m+1}\right)^2 \tag{3.94}$$

for $l = 1, \ldots, N$. However, when demand information is not provided,

$$\frac{VAR(x_t^l)}{\sigma^2} > \prod_{u=1}^{l}\left[1 + 2\frac{\tau_u}{m+1} + 2\left(\frac{\tau_u}{m+1}\right)^2\right] \ , l = 1, \ldots, N. \tag{3.95}$$

In the first situation the ratio increases additively.

From echelon l to $l+1$,

$$\frac{VAR(x_t^{l+1})}{VAR(x_t^l)} = 1 + \frac{\left(\frac{2\tau_{l+1}}{m+1} + \frac{2(\tau_{l+1}^2 + 2\tau_{l+1}\sum_{u=1}^l \tau_u)}{(m+1)^2}\right)}{\left(1 + \frac{2\sum_{u=1}^l \tau_u}{m+1} + 2(\frac{\sum_{u=1}^l \tau_u}{m+1})^2\right)} .$$

However, when demand information is not passed throughout the system, the increase in the lower bound for the ratio at each echelon is multiplicative, that is

$$\frac{VAR(x_t^l)}{\sigma^2} \geq \frac{VAR(x_t^{l-1})}{\sigma^2} \cdot \left[1 + 2\frac{\tau_l}{m+1} + 2\left(\frac{\tau_l}{m+1}\right)^2\right]. \tag{3.96}$$

Thus by sharing the information, the variability of the order process may be reduced dramatically. However, as we have mentioned previously, even with information sharing, the bullwhip effect can not be eliminated. The effect of reducing lead times on order quantity variability is also evident and obvious to managers of supply chains.

4 The Two-Stage System Revisited

In the second section we examined a two-stage or echelon system and determined how the variance of the order size compared with the variance of the demand process at the DC when a moving average method was used to forecast future demand. Our goal in this section is to analyze the variance of the order size when a different forecast method is used, an exponential smoothing method. We will conclude this section by comparing the effects of using these two forecasting methods on order quantity variance.

4.1 Determining Order Quantity Variance

When the DC uses an exponential smoothing forecast to estimate the average period demand μ, the estimate is computed as

$$F_t^1 = \alpha d_t + (1-\alpha)F_{t-1}^1, \tag{4.97}$$

where α is the smoothing constant and $\alpha \in (0,1)$. Thus F_t^1 is a weighted average of the period t demand, d_t, and the previous period's forecast, F_{t-1}^1. Assuming the system has been operating for an infinite number of periods, we see that

$$F_t^1 = \alpha \sum_{j=0}^{\infty} (1-\alpha)^j d_{t-j}. \tag{4.98}$$

Note that

$$E[F_t^1] = E[\alpha \sum_{j=0}^{\infty} (1-\alpha)^j d_{t-j}] \tag{4.99}$$

$$= \alpha \sum_{j=0}^{\infty} (1-\alpha)^j E[d_{t-j}] \tag{4.100}$$

$$= \mu, \tag{4.101}$$

that is, F_t^1 is an unbiased estimator of the mean demand per period. Then $F_t^\tau = \tau \cdot F_t^1$ measures the expected lead time demand.

As in section 2, we estimate the τ period forecast error (the sample variance). We could also use the mean absolute deviation (MAD) of the forecast errors to estimate the process standard deviation. This is what is frequently done in practice. (When a MAD method is used, it is commonly assumed that the forecast errors are normally distributed.)

Now that we have stated how we will estimate the mean and standard deviation of demand over a lead time, we can turn our attention to the ordering policy. As we assumed earlier, the DC uses an order-up-to policy in each period. The order-up-to level in period t is given by

$$y_t = F_t^\tau + k\hat{S}_t^\tau. \tag{4.102}$$

The order size in period t is given by

$$\begin{aligned} x_t &= y_t - y_{t-1} + d_t && (4.103) \\ &= (F_t^\tau - F_{t-1}^\tau) + k(\hat{S}_t^\tau - \hat{S}_{t-1}^\tau) + d_t && (4.104) \end{aligned}$$

We are interested in knowing the variance of x_t, which is computed as follows.

$$\begin{aligned} VAR(x_t) &= VAR[F_t^\tau - F_{t-1}^\tau + d_t] + k^2 VAR[\hat{S}_t^\tau - \hat{S}_{t-1}^\tau] && (4.105) \\ &\quad +2 \cdot COV(F_t^\tau - F_{t-1}^\tau + d_t, k(\hat{S}_t^\tau - \hat{S}_{t-1}^\tau)). && (4.106) \end{aligned}$$

If we further assume that the demand process is i.i.d and symmetric about μ, then we can show that the covariance term in the preceding equation is zero, that is, in this case

$$\begin{aligned} VAR(x_t) &= VAR[F_t^\tau - F_{t-1}^\tau + d_t] + k^2 VAR[\hat{S}_t^\tau - \hat{S}_{t-1}^\tau] && (4.107) \\ \text{and} & && (4.108) \\ VAR(x_t) &\geq VAR[F_t^\tau - F_{t-1}^\tau + d_t] && (4.109) \end{aligned}$$

$$= VAR[(1+\tau\alpha)d_t + (1-\alpha)F^{\tau}_{t-1} - F^{\tau}_{t-1}] \tag{4.110}$$

$$= VAR[(1+\tau\alpha)d_t - \alpha F^{\tau}_{t-1}] \tag{4.111}$$

$$= (1+\alpha\tau)^2\sigma^2 + \alpha^2\tau^2 VAR(F^1_{t-1}). \tag{4.112}$$

But

$$VAR(F^1_{t-1}) = VAR[\alpha\sum_{j=0}^{\infty}(1-\alpha)^j d_{t-j-1}] \tag{4.113}$$

$$= \alpha^2\sum_{j=0}^{\infty} VAR[(1-\alpha)^j d_{t-j-1}] \tag{4.114}$$

$$= \alpha^2\sum_{j=0}^{\infty}((1-\alpha)^j)^2\sigma^2 \tag{4.115}$$

$$= \alpha^2\sigma^2\sum_{j=0}^{\infty}(1-\alpha)^{2j} \tag{4.116}$$

$$= \alpha^2\sigma^2\sum_{j=0}^{\infty}(1-2\alpha+\alpha^2)^j\ . \tag{4.117}$$

Let $\beta = 1-2\alpha+\alpha^2$, then we have

$$VAR(F^1_{t-1}) = \alpha^2\sigma^2\sum_{j=0}^{\infty}\beta^j \tag{4.118}$$

$$= \frac{\alpha^2\sigma^2}{1-\beta}\cdot\sum_{j=0}^{\infty}\beta^j(1-\beta) \tag{4.119}$$

$$= \frac{\alpha^2\sigma^2}{2\alpha-\alpha^2} = \frac{\alpha\sigma^2}{2-\alpha}. \tag{4.120}$$

Thus,

$$\frac{VAR(x_t)}{\sigma^2} \geq 1+2\alpha\tau+(\alpha\tau)^2\left(1+\frac{\alpha}{2-\alpha}\right) \tag{4.121}$$

$$= 1+2\alpha\tau+(\alpha\tau)^2\frac{2}{2-\alpha}. \tag{4.122}$$

If the inventory policy is to set $k=0$, then the above inequality is a strict equality.

4.2 A Comparison of the Moving Average and Exponential Smoothing Methods

Let us now compare how the variability in the order quantity differs between the two forecasting methods that we have discussed. To begin, suppose the values of m and α have been chosen so that the true forecast errors are the same. The variance of the single period forecast error is

$$s^2 = \frac{m+2}{m+1}\sigma^2 \tag{4.123}$$

$$\text{or } s = \sqrt{\frac{m+2}{m+1}}\sigma \tag{4.124}$$

when the moving average method is used to forecast demand.

When the exponential smoothing method is used,

$$s^2 = \frac{2}{2-\alpha}\sigma^2 \tag{4.125}$$

$$\text{or } s = \sqrt{\frac{2}{2-\alpha}}\sigma. \tag{4.126}$$

If we equate the variances of these two estimators, we see that we must set

$$\alpha = \frac{2}{m+2}. \tag{4.127}$$

Suppose $k = 0$. Then from 4.122 we know that

$$\frac{VAR(x_t)}{\sigma^2}(\text{Exp. Case}) = 1 + 2\alpha\tau + \frac{2(\alpha\tau)^2}{2-\alpha} \tag{4.128}$$

$$= 1 + \frac{4\tau}{m+2} + \frac{2\frac{4\tau^2}{(m+2)^2}}{2-\frac{2}{m+2}} \tag{4.129}$$

$$= 1 + \frac{4\tau}{m+2} + \frac{4\tau^2}{(m+1)(m+2)}. \tag{4.130}$$

But, in the moving average case,

$$\frac{VAR(x_t)}{\sigma^2}(\text{Mov. avg. case}) = 1 + \frac{2\tau}{m+1} + \frac{2\tau^2}{(m+1)^2}. \tag{4.131}$$

Observe that

$$
\begin{aligned}
1+\frac{4\tau}{m+2}+\frac{4\tau^2}{(m+1)(m+2)} &= 1+\frac{1}{m+2}\left[4\tau+\frac{4\tau^2}{m+1}\right] && (4.132)\\
&= 1+\frac{2(m+1)}{m+2}\left[\frac{2\tau}{m+1}+2\left(\frac{\tau}{m+1}\right)^2\right] && (4.133)\\
&> 1+\frac{2\tau}{m+1}+2\left(\frac{\tau}{m+1}\right)^2, \text{ for } m>0. && (4.134)
\end{aligned}
$$

Hence the variability in the order size is greater when an exponential smoothing forecasting method is used. Thus even if α and m are chosen to give equal forecast errors, the order size variance will be greater when an exponential smoothing method is used to forecast demand.