Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times and Information*

Frank Chen
Decision Sciences Dept., National University of Singapore
Zvi Drezner
Dept. of MS & IS, California State University, Fullerton
Jennifer K. Ryan
School of Industrial Engineering, Purdue University
David Simchi-Levi
Dept. of IE & MS, Northwestern University

Submitted: October 1996 Revised: December 1998, June 14, 1999

Abstract

An important observation in supply chain management, known as the **bullwhip effect**, suggests that demand variability increases as one moves up a supply chain. In this paper we quantify this effect for simple, two-stage, supply chains consisting of a single retailer and a single manufacturer. Our model includes two of the factors commonly assumed to cause the bullwhip effect: demand forecasting and order lead times. We extend these results to multiple stage supply chains with and without centralized customer demand information and demonstrate that the bullwhip effect can be reduced, but not completely eliminated, by centralizing demand information.

1 Introduction

An important observation in supply chain management, known as the **bullwhip effect**, suggests that demand variability increases as one moves up a supply chain. For a detailed discussion of this phenomenon see Baganha and Cohen (1995), Kahn (1987), Lee, Padmanabhan and Whang (1997a, b), and Metters (1996).

Most of the previous research on the bullwhip effect has focused on demonstrating its existence, identifying its possible causes, and providing methods for reducing its impact. In particular, Lee, Padmanabhan and Whang (1997a, b) identify five main causes of the bullwhip effect: the use

^{*}Research supported in part by ONR Contracts N00014-90-J-1649 and N00014-95-1-0232, NSF Contracts DDM-8922712, DDM-9322828 and DMI-9732795.

of demand forecasting, supply shortages, lead times, batch ordering, and price variations. This previous work has also led to a number of approaches and suggestions for reducing the impact of the bullwhip effect. For instance, one of the most frequent suggestions is the centralization of demand information, that is, providing each stage of the supply chain with complete information on customer demand.

This paper differs from the previous research in several ways. First, our focus is on determining the impact of **demand forecasting** on the bullwhip effect. In other words, we do not assume that the retailer knows the exact form of the customer demand process. Instead, the retailer uses a standard forecasting technique to estimate certain parameters of the demand process. Second, the goal is not only to demonstrate the existence of the bullwhip effect, but also to quantify it, i.e., to quantify the increase in variability at each stage of the supply chain.

To quantify the increase in variability from the retailer to the manufacturer we first consider a **simple two-stage** supply chain consisting of a single retailer and a single manufacturer. In Section 2 we describe the supply chain model, the forecasting technique and the inventory policy used by the retailer. We derive a lower bound for the variance of the orders placed by the retailer to the manufacturer. In Section 3 we consider a multi-stage supply chain and the impact of centralized demand information on the bullwhip effect. Finally, in Section 4, we conclude with a discussion of our results.

2 A Simple Supply Chain Model

Consider a simple supply chain in which in each period, t, a single retailer observes his inventory level and places an order, q_t , to a single manufacturer. After the order is placed, the retailer observes and fills customer demand for that period, denoted by D_t . Any unfilled demands are backlogged. There is a fixed lead time between the time an order is placed by the retailer and when it is received at the retailer, such that an order placed at the end of period t is received at the start of period t + L. For example, if there is no lead time, an order placed at the end of period t is received at the start of period t + 1 and thus t = 1.

The customer demands seen by the retailer are random variables of the form

$$D_t = \mu + \rho D_{t-1} + \epsilon_t \tag{1}$$

where μ is a non-negative constant, ρ is a correlation parameter with $|\rho| < 1$, and the error terms, ϵ_t , are independent and identically distributed (i.i.d.) from a symmetric distribution with mean 0 and variance σ^2 . It can easily be shown that $E(D_t) = \frac{\mu}{1-\rho}$ and $Var(D_t) = \frac{\sigma^2}{1-\rho^2}$. Demand processes of this form have been assumed by several authors analyzing the bullwhip effect (e.g., Lee, et al. and Kahn). Finally, note that if $\rho = 0$, equation (1) implies that the demands are i.i.d. with mean μ and variance σ^2 .

2.1 The Inventory Policy and Forecasting Technique

We assume that the retailer follows a simple **order-up-to inventory policy** in which the orderup-to point, y_t , is estimated from the observed demand as

$$y_t = \hat{D}_t^L + z\hat{\sigma}_{et}^L,\tag{2}$$

where \hat{D}_t^L is an estimate of the mean lead time demand, $\hat{\sigma}_{et}^L$ is an estimate of the standard deviation of the L period forecast error, and z is a constant chosen to meet a desired service level. It is well known that for i.i.d. demands from a normal distribution an order-up-to policy of this form is optimal.

Note that the order-up-to point is calculated based on the standard deviation of the L period forecast error, σ_e^L , whose estimator is $\hat{\sigma}_{et}^L$, and not based on the standard deviation of the lead time demand, σ^L , whose estimator is $\hat{\sigma}_t^L$. While there is a simple relationship between these two quantities, i.e., $\sigma_e^L = c\sigma^L$, for some constant $c \geq 1$, it is more appropriate to calculate the inventory policy based on the former quantity (Hax and Candea, 1984, p. 194).

We assume that the retailer uses a simple moving average to estimate \hat{D}_t^L and $\hat{\sigma}_{et}^L$ based on the demand observations from the previous p periods. That is,

$$\hat{D}_t^L = L\left(\frac{\sum_{i=1}^p D_{t-i}}{p}\right)$$

and

$$\hat{\sigma}_{et}^{L} = C_{L,\rho} \sqrt{\frac{\sum_{i=1}^{p} (e_{t-i})^2}{p}},$$
(3)

where e_t is the one period forecast error, i.e., $e_t = D_t - \hat{D}_t^1$, and $C_{L,\rho}$ is a constant function of L, ρ and ρ . See Ryan (1997) for a more detailed discussion of this constant.

2.2 Quantifying the Bullwhip Effect

Our objective is to quantify the bullwhip effect. To do this, we must determine the variance of q_t relative to the variance of D_t , i.e., the variance of the orders placed by the retailer to the manufacturer relative to the variance of the demand faced by the retailer. For this purpose, we write q_t as

$$q_t = y_t - y_{t-1} + D_{t-1}$$
.

Observe that q_t may be negative, in which case we assume, similarly to Kahn (1987) and Lee, Padmanabhan and Whang (1997b), that this excess inventory is returned without cost. We discuss the impact of this assumption on our results later in this section.

Given the estimates of the lead time demand, \hat{D}_t^L , and the standard deviation of the L period forecast error, $\hat{\sigma}_{et}^L$, we can write the order quantity q_t as

$$q_{t} = \hat{D}_{t}^{L} - \hat{D}_{t-1}^{L} + z(\hat{\sigma}_{et}^{L} - \hat{\sigma}_{e,t-1}^{L}) + D_{t-1}$$

$$= L\left(\frac{D_{t-1} - D_{t-p-1}}{p}\right) + D_{t-1} + z(\hat{\sigma}_{et}^{L} - \hat{\sigma}_{e,t-1}^{L})$$

$$= (1 + L/p)D_{t-1} - (L/p)D_{t-p-1} + z(\hat{\sigma}_{et}^{L} - \hat{\sigma}_{e,t-1}^{L}). \tag{4}$$

It follows that

$$\begin{aligned} Var(q_t) &= (1 + L/p)^2 Var(D_{t-1}) - 2(L/p)(1 + L/p)Cov(D_{t-1}, D_{t-p-1}) + (L/p)^2 Var(D_{t-p-1}) \\ &+ z^2 Var(\hat{\sigma}_{et}^L - \hat{\sigma}_{e,t-1}^L) + 2z(1 + 2L/p)Cov(D_{t-1}, \hat{\sigma}_{et}^L) \\ &= \left(1 + \frac{2L}{p} + \frac{2L^2}{p^2}\right) Var(D) - \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) \rho^p Var(D) \\ &+ 2z(1 + 2L/p)Cov(D_{t-1}, \hat{\sigma}_{et}^L) + z^2 Var(\hat{\sigma}_{et}^L - \hat{\sigma}_{e,t-1}^L) \\ &= \left[1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)(1 - \rho^p)\right] Var(D) \\ &+ 2z(1 + 2L/p)Cov(D_{t-1}, \hat{\sigma}_{et}^L) + z^2 Var(\hat{\sigma}_{et}^L - \hat{\sigma}_{e,t-1}^L), \end{aligned}$$

where the second equation follows from $Cov(D_{t-1}, D_{t-p-1}) = \frac{\rho^p}{(1-\rho^2)}\sigma^2$ and $Var(D_t) = \frac{\sigma^2}{1-\rho^2}$.

To further evaluate $Var(q_t)$ we need the following lemma:

Lemma 2.1 Assume the customer demands seen by the retailer are random variables of the form given in (1) where the error terms, ϵ_t , are i.i.d. from a symmetric distribution with mean 0 and variance σ^2 . Let the estimate of the standard deviation of the L period forecast error be $\hat{\sigma}_{et}^L$, as defined in (3). Then

$$Cov(D_{t-i}, \hat{\sigma}_{et}^L) = 0$$
 for all $i = 1, \dots, p$.

The interested reader is referred to Ryan (1997) or Chen, Drezner, Ryan and Simchi-Levi (1998) for the proof of Lemma 2.1.

We thus have the following lower bound on the increase in variability from the retailer to the manufacturer:

Theorem 2.2 Under the conditions of Lemma 2.1, if the retailer uses a simple moving average forecast with p demand observations, then the variance of the orders, q, placed by the retailer to the manufacturer, satisfies

$$\frac{Var(q)}{Var(D)} \ge 1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)(1 - \rho^p). \tag{5}$$

The bound is tight when z = 0.

2.3 Interpretation of the Results

Figure 1 shows the lower bound (solid lines) on the increase in variability for L=1 and z=2 for various values of ρ and p. It also presents simulation results on $Var(q_t)/\sigma^2$ (dotted lines), as well as the increase in variability when excess inventory is not returned (dashed lines). That is, the dashed lines represent simulation results on $Var(q_t^+)/\sigma^2$, where

$$q_t^+ = \max\{q_t, 0\} = \max\{y_t - y_{t-1} + D_{t-1}, 0\}.$$

Notice that, given ρ , Var(q) and $Var(q^+)$ are quite close. This implies that **the way one treats** excess inventory has little impact on the increase in variability for most values of p and ρ . Although this observation applies for all positive values of ρ , we should note that when $\rho \to -1$, Var(q) can be greater than $Var(q^+)$.

Figure 2 shows the lower bound (solid lines) on the variability amplification when $\rho = 0$ as a function of p for various values of the lead time parameter, L. It also presents simulation results on $Var(q_t)/\sigma^2$ (dotted lines), as well as the increase in variability when excess inventory is not returned (dashed lines). That is, the dashed lines represent simulation results on $Var(q_t^+)/\sigma^2$. Notice that, given L, the three lines are quite close. This, again, implies that the way one treats excess inventory has little impact on the increase in variability.

Several observations regarding the increase in variability can be made from (5). First, we notice that the increase in variability from the retailer to the manufacturer is a function of three parameters, (1) p, the number of observations used in the moving average, (2) L, the lead time parameter, and (3) ρ , the correlation parameter.

More specifically, the increase in the variability of orders from the retailer to the manufacturer is a decreasing function of p, the number observations used to estimate the mean and variance of demand. In particular, when p is large the increase in variability is negligible. However, when p is small, there can be a significant increase in variability. In other words, the smoother the demand forecasts, the smaller the increase in variability.

Also note that the increase in the variability of orders from the retailer to the manufacturer is an increasing function of L, the lead time parameter. In addition, note the relationship between L and p: if the lead time parameter doubles, then we must use twice as much demand data in order to maintain the same variability in the order process. In other words, with longer lead times, the retailer must use more demand data in order to reduce the bullwhip effect.

The correlation parameter, ρ , can also have a significant impact on the increase in variability. First, if $\rho = 0$, i.e., if demands are i.i.d., then

$$\frac{Var(q)}{Var(D)} \ge 1 + \frac{2L}{p} + \frac{2L^2}{p^2}.$$
 (6)

Second, if $\rho \geq 0$, i.e., demands are positively correlated, $(1 - \rho^p) < 1$, and the larger ρ , the smaller the increase in variability. This can also be seen in Figure 1. On the other hand, if $\rho < 0$, i.e., demands are negatively correlated, then we see some strange behavior. For even values of p, $(1 - \rho^p) < 1$, while for odd values of p, $(1 - \rho^p) > 1$. Therefore, for $\rho < 0$, the lower bound on the increase in variability will be larger for odd values of p than for even values of p.

3 The Impact of Centralized Demand Information

One frequently suggested strategy for reducing the magnitude of the bullwhip effect is to centralize demand information, i.e., to make customer demand information available to every stage of the supply chain. For example, Lee, Padmanabhan, and Whang (1997a) suggest that "one remedy is to make demand data at a downstream site available to the upstream site." In this section, we analyze the impact of centralized customer demand information on the bullwhip effect. In particular, we demonstrate that while it will certainly reduce the magnitude of the bullwhip effect, centralizing demand information will not completely eliminate the increase in variability. That is, even if each stage of the supply chain has complete knowledge of the demands seen by the retailer, the bullwhip effect will still exist.

Consider a multi-stage supply chain in which the first stage (i.e., the retailer) shares all demand information with each of the subsequent stages. Assume that all stages in the supply chain use a moving average forecast with p observations to estimate the mean demand. Therefore, since each stage has complete information on customer demand, each stage will use the same estimate of the mean demand per period, $\hat{D}_t = \frac{\sum_{i=1}^p D_{t-i}}{p}$.

Assume that each stage k follows an order-up-to inventory policy where the order-up-to point is of the form

$$y_t^k = L_k \hat{D}_t + z_k \hat{\sigma}_{et}^{L_k},$$

where \hat{D}_t is the estimate of the mean demand per period, L_k is the lead time between stages k and k+1, $\hat{\sigma}_{et}^{L_k} = C_{L_k,\rho} \sqrt{\frac{\sum_{i=1}^p (e_{t-i})^2}{p}}$, and z_k is a constant.

We note that the assumption that all stages in the supply chain use the same demand data, the same inventory policy and the same forecasting technique allows us to determine the impact of just the demand forecasting, without considering the impact of different forecasting techniques or different inventory policies across the stages.

The sequence of events in our model is similar to the one in the traditional Beer Game (Sterman 1989) or the computerized Beer Game (Kaminsky and Simchi-Levi 1998). At the end of period t-1, the retailer, or stage 1, observes customer demand, D_{t-1} , calculates his order-up-to point for period t, y_t^1 , and orders q_t^1 so as to raise his inventory to level y_t^1 . The manufacturer, or stage 2, receives the order, q_t^2 , along with the most recent demand information, D_{t-1} . This information is

received at the end of period t-1 (i.e., there is no information lead time), allowing the manufacturer to calculate his order-up-to point, y_t^2 , and immediately place an order of q_t^2 in order to raise his inventory to level y_t^2 . The next stage, stage 3, immediately receives the order, q_t^2 , and the demand observation D_{t-1} , and the process continues.

We can now state the following lower bound on the variance of the orders placed by each stage of the supply chain relative to the variance of customer demand.

Theorem 3.1 Consider a multi-stage supply chain where the demands seen by the retailer are random variables of the form given in (1) where the error terms, ϵ_t , are i.i.d. from a symmetric distribution with mean 0 and variance σ^2 . Each stage of the supply chain follows an order-up-to policy of the form $y_t^k = L_k \hat{D}_t + z_k \hat{\sigma}_{et}^{L_k}$, where L_k is the lead time between stages k and k + 1.

The variance of the orders placed by stage k, denoted q^k , satisfies

$$\frac{Var(q^k)}{Var(D)} \ge 1 + \left(\frac{2(\sum_{i=1}^k L_i)}{p} + \frac{2(\sum_{i=1}^k L_i)^2}{p^2}\right) (1 - \rho^p) \qquad \forall k$$

This bound is tight when $z_i = 0$ for i = 1, ..., k.

The reader is referred to Ryan (1997) or Chen, Drezner, Ryan and Simchi-Levi (1998) for a proof of Theorem 3.1.

This result demonstrates that even when (i) all demand information is centralized, (ii) every stage of the supply chain uses the same forecasting technique, and (iii) every stage uses the same inventory policy, there will still be an increase in variability at every stage of the supply chain. In other words, we have not completely eliminated the bullwhip effect by centralizing customer demand information.

Finally, it would be interesting to compare the increase in variability for the case of centralized information, as given in Theorm 3.1, with the increase in the case of decentralized information. To do this, we consider a supply chain similar to the one just analyzed, but without centralized information. In this case, the retailer does not provide the upstream stages with any customer demand information and each stage determines its forecast demand based on the orders placed by the previous stage, not based on actual customer demands. In this case, we can state the following lower bound on the variance of the orders placed by each stage of the supply chain.

Theorem 3.2 Consider a multi-stage supply chain where the demands seen by the retailer are random variables of the form $D_t = \mu + \epsilon_t$, where the error terms, ϵ_t , are i.i.d. from a symmetric distribution with mean 0 and variance σ^2 . Each stage of the supply chain follows an order-up-to policy of the form $y_t^k = L_k \hat{D}_t^{(k)}$, where L_k is the lead time between stages k and k + 1,

$$\hat{D}_{t}^{(1)} = \frac{\sum_{i=1}^{p} D_{t-i}}{p} \quad and \quad \hat{D}_{t}^{(k)} = \frac{\sum_{j=0}^{p-1} q_{t-j}^{k-1}}{p} \quad for \ k \ge 2,$$

where q_t^k is the order placed by stage k in period t.

The variance of the orders placed by stage k, denoted q^k , satisfies

$$\frac{Var(q^k)}{Var(D)} \ge \prod_{i=1}^k \left(1 + \frac{2L_i}{p} + \frac{2L_i^2}{p^2} \right) \qquad \forall \ k.$$
 (7)

The reader is referred to Ryan (1997) or Chen, Drezner, Ryan and Simchi-Levi (1998) for a proof of Theorem 3.2.

Note that Theorem 3.2 only applies when each stage of the supply chain uses an inventory policy as defined in (2) with z = 0. When a policy of this form is used in practice, an inflated value of L_k is often used, with the excess inventory representing the safety stock. For example, a retailer facing an order lead time of three weeks may choose to keep inventory equal to four weeks of forecast demand, with the extra week of inventory representing his safety stock. Policies of this form are often used in practice. Indeed, we have recently collaborated with a major U.S. retail company that uses a policy of this form. See also Johnson *et al.* (1995).

We are interested in comparing the increase in variability at each stage of the supply chain for the centralized and decentralized systems. We'll consider the case of i.i.d. demands, i.e., $\rho = 0$, and $z_i = 0$ for all i. For a discussion of the case where $\rho \neq 0$, see Ryan (1997) or Chen, Drezner, Ryan and Simchi-Levi (1998). If demand information is shared with each stage of the supply chain, then the increase in variability from the retailer to stage k is

$$\frac{Var(q^k)}{Var(D)} = 1 + \frac{2(\sum_{i=1}^k L_i)}{p} + \frac{2(\sum_{i=1}^k L_i)^2}{p^2}$$
(8)

On the other hand, if demand information is not shared with each stage of the supply chain, then a lower bound on the increase in variability from the retailer to stage k is given by (7).

Note that for supply chains with centralized information, the increase in variability at each stage is an additive function of the lead time and the lead time squared, while for supply chains without centralized information, the lower bound on the increase in variability at each stage is multiplicative. This implies that centralizing customer demand information can significantly reduce the bullwhip effect. However, as mentioned above, centralizing customer demand information does not completely eliminate the bullwhip effect. In addition, these results indicate that the difference between the variability in the centralized and decentralized supply chains increases as we move up the supply chain, i.e., as we move from the first stage to the second and third stages of the supply chain. A simulation study has confirmed these insights for the case in which $z_i \neq 0$ for all i, and $\rho \neq 0$. See see Ryan (1997) or Chen, Drezner, Ryan and Simchi-Levi (1998) for the results of this simulation study.

4 Summary

In this paper we have demonstrated that the phenomenon known as the bullwhip effect is due, in part, to the effects of demand forecasting. In particular, we have shown that if a retailer periodically updates the mean and variance of demand based on observed customer demand data, then the variance of the orders placed by the retailer will be greater than the variance of demand. More importantly, we have shown that providing each stage of the supply chain with complete access to customer demand information can significantly reduce this increase in variability. However, we have also shown that the bullwhip effect will exist even when demand information is shared by all stages of the supply chain and all stages use the same forecasting technique and inventory policy.

To further explain this point, consider the first stage of the supply chain, the retailer. Notice that if the retailer knew the mean and the variance of customer demand, then the orders placed by the retailer would be exactly equal to the customer demand and there would be no increase in variability. In our model, however, even though the retailer has complete knowledge of the observed customer demands, he must still estimate the mean and variance of demand. As a result, the manufacturer sees an increase in variability.

This paper would be incomplete if we do not mention several important limitations of our model and results. The main drawback of our model is the assumption that excess inventory is returned without cost. The simulation results demonstrate, however, that this assumption has little impact

on the increase in variability for all reasonable values of p, ρ and L. Another limitation of our model is the fact that we are not using the optimal order-up-to policy, but rather the policy defined by equation (2). It is important to point out, however, that policies of this form are frequently used in practice. In addition, we are not using the optimal forecasting technique for the correlated demand process considered here. However, we again point out that the moving average is one of the most commonly used forecasting techniques in practice. Indeed, we believe that when evaluating the bullwhip effect it is most appropriate to consider inventory policies and forecasting techniques that are used in practice.

Finally, our model clearly does not capture many of the complexities involved in real-world supply chains. For example, we have not considered a multi-stage system with multiple retailers and manufacturers. Fortunately, extending our results to the multi-retailer case when demand between retailers may be correlated is straightforward. See Ryan (1997).

5 References

- Baganha, M. and M. Cohen. (1995), The Stabilizing Effect of Inventory in Supply Chains. To appear in *Operations Research*.
- Chen, F., Drezner, Z., Ryan J. and Simchi-Levi, D. (1998), Quantifying the Bullwhip Effect: The Impact of Forecasting, Lead Times and Information. Working Paper, Department of IE/MS Northwestern University and School of Industrial Engineering, Purdue University.
- Kahn, J. (1987), Inventories and the Volatility of Production. The American Economic Review, 77, pp. 667–679.
- Kaminsky, P. and Simchi-Levi, D. (1998) A New Computerized Beer Game: A Tool for Teaching the Value of Integrated Supply Chain Management. **Global Supply Chain and Technology Management**. Hau Lee and Shu Ming Ng, eds., POMS Series in Technology and Operations Management, pp. 216–225.
- Lee, H., Padmanabhan, P. and Whang, S. (1997a), The Bullwhip Effect in Supply Chains. *Sloan Management Review*, **38**, pp. 93-102.
- Lee, H., Padmanabhan, P. and Whang, S. (1997b), Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Science*, **43**, pp. 546–58.
- Metters, R. (1996), Quantifying the Bullwhip Effect in Supply Chains. *Proceedings of the 1996 MSOM Conference*, pp. 264–269.
- Porteus, E. (1990), Stochastic Inventory Theory. Handbooks in Operations Research and Management Science, Volume 2, pp. 605-651.
- Ryan, J. K. (1997), Analysis of Inventory Models with Limited Demand Information. Ph.D. Dissertation. Department of Industrial Engineering and Management Science, Northwestern University.
- Sterman, J. D. (1989), Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science*, **35**, pp. 321-339.