

**BEHAVIORAL CAUSES OF THE BULLWHIP EFFECT
AND THE OBSERVED VALUE OF INVENTORY INFORMATION**

Rachel Croson

Department of Operations and Information Management
The Wharton School, University of Pennsylvania
Philadelphia, PA 19104-6366
crosonr@wharton.upenn.edu

Karen Donohue

Department of Operations and Management Science
The Carlson School, University of Minnesota
Minneapolis, MN 55455-9940
kdonohue@csom.umn.edu

Abstract

The tendency of orders to increase in variability as one moves up a supply chain is commonly known as the ‘bullwhip effect.’ We study this phenomenon from a behavioral perspective in the context of a simple, serial, supply chain subject to information lags and stochastic demand. We conduct two experiments on two different sets of participants. In the first we find the bullwhip effect still exists when normal operational causes (e.g., batching, price fluctuations, demand estimation, etc.) are removed. The persistence of the bullwhip effect is explained, to some extent, by evidence that decision-makers consistently under-weight the supply line when making order decisions. In the second experiment we find that the bullwhip, and the underlying tendency of underweighting, remains when information on inventory levels is shared. However, we observe that inventory information helps somewhat to alleviate the bullwhip effect by helping upstream chain members better anticipate and prepare for fluctuations in inventory needs downstream. These experimental results support the theoretically suggested notion that upstream chain members stand to gain the most from information sharing initiatives.

The first author’s research was supported by NSF Grant #SBR 9753130. The second author’s research was supported by NSF Career Award #SBR 9602071.

1 Introduction

In recent years there has been a vast increase in the quality and quantity of information shared across supply chains. This increase is driven, in part, by improvements in the technology available for gathering and sharing data. The advent of enterprise logistics software, such as SAP, now allows companies to maintain and share inventory information for multiple supply points on a common database (Stein 1998). Many industry analysts claim that such systems offer tremendous savings for integrated supply chains, despite their significant investment cost. In the auto industry, for example, analysts project that improvements in the quality of information shared between OEMs up through the highest level in the supply chain will save \$1 billion in supply-chain inefficiencies over the next few years alone (Scheck 1998). One explanation given for why information sharing leads to lower costs is that it may reduce the ‘bullwhip effect.’

The bullwhip effect refers to the tendency of orders to increase in variation as one moves up a supply chain. Forrester (1958) was the first to point out this effect and its possible causes. Increased variation is a concern for distribution chains since it leads to increased costs in the form of increased inventory requirements, expediting, or customer shortages. In the last few years, supply chain managers, as well as academics, have focused attention on the *operational* causes of the bullwhip effect. These causes include demand signal processing, inventory rationing, order batching, and price variations (Lee et al. 1997a and 1997b). Ways to alleviate these operational problems include improved demand forecasting techniques (Chen et al. 1998), capacity allocation schemes (Cachon and Lariviere 1999), staggered order batching (Cachon 1999), and everyday low pricing (Sogomonian and Tang 1993).

The goals of this paper, in contrast, are to shed light on the *behavioral* causes of the bullwhip and investigate the potential benefit inventory information sharing offers for

overcoming these problems. We report on the results of two controlled experiments, each patterned after the well-known ‘Beer Distribution Game’ (Sterman 1992), where players make ordering decisions within a simple, linear, supply chain subject to information and shipment lags. Our first experiment reveals that the bullwhip effect still exists when traditional operational causes are removed and the demand distribution is known by all parties. This seemingly ‘irrational’ behavior is in line with previous behavioral research (e.g., Kahneman and Lovallo 1993, Schotter and Weigelt 1992, Schweitzer and Cachon 2000), which shows that individuals often exhibit some form of decision bias in business settings. We test for the existence of one particular type of decision bias, under-weighting the supply line, which was first identified by Sterman (1989a) in his study of the bullwhip effect. Results from our first study confirm that decision-makers still exhibit this form of decision bias in a controlled setting where demand information is known and stationary.

In our second study we examine the behavioral impact of inventory tracking systems, which expose the inventory position of each member of the supply chain and therefore may help to overcome this under-weighting bias. We find that exposure to real time inventory information helps reduce the bullwhip effect, but not in the manner expected. Decision-makers continue to under-weight the supply line, even though stock is now in full view. As a result, exposure to such information has little impact on the variance of orders of downstream chain members. On the other hand, this information appears to substantially reduce the variance of orders for upstream members. We postulate that exposure to inventory information helps these chain members better prepare for fluctuations in orders downstream.

This observation has important implications for companies looking to invest in inventory tracking systems. From a behavioral point of view, our results suggest that upstream members stand to gain the most from such an investment. The results also suggest that humans are still

prone to decision biases when such an investment is made, even in this simplified supply chain setting. As one anonymous referee points out, this gives further credence to the use of automated ordering systems, such as those often bundled with inventory tracking software.

The paper continues in section 2 with a review of relevant literature, followed by a description of our experimental setting in section 3. Section 4 reports on the results of our first experiment, which demonstrates the existence of the bullwhip effect while controlling for operational causes. Section 5 describes the results of our second experiment, which focuses on the impact of sharing inventory information. Section 6 provides final conclusions and a framework for future research.

2 Literature Review

Sterman (1989a) was the first to use the Beer Distribution Game to rigorously test the existence of the bullwhip effect in an experimental context. His experiments rely on a simple, non-stationary retail demand function (beginning at four cases per period and jumping to eight) unknown to chain members. In this sense, his experiments control for three out of the four operational causes cited by Lee *et al.* (1997b); inventory rationing, order batching, and price fluctuations. However, demand signal processing still looms as an operational issue since the demand distribution is non-stationary and unknown. Within this environment, Sterman provides evidence that the bullwhip effect exists and may be caused by participants' tendency to under-weight inventory in the supply line. That is, participants place orders in one period, but do not account for them in their inventory calculations when placing orders in the next period. Thus inventory that has been ordered but not yet received (i.e., is in the supply line) is under-weighted in the decision to order more. We build on this research by showing, in our first study, that both

the bullwhip effect and under-weighting still occur when all operational causes are removed (i.e., when the demand distribution is stable and known to chain members).

Subsequent researchers have used the Beer Distribution Game to test various strategies for reducing the bullwhip effect (e.g., Kaminsky and Simchi-Levi 1998, Gupta et al. 1998). Common strategies include reducing order lead-times, sharing POS data, and centralizing ordering decisions. Anderson and Morrice (2000) study a novel service-oriented supply chain game in a similar manner, showing that sharing consumer demand (in the form of customer orders feeding into a make-to-order Mortgage Service) can lead to improvements in capacity fluctuations and cost. Our research is the first to examine the behavioral impact of sharing echelon inventory information in a supply chain. Our research also differs from previous experimental studies in its level of control. We use a computerized version of the game, like Kaminsky and Simchi-Levi (1998) and Gupta et al. (1998), that avoids accounting errors and carefully controls what information is available to each participant and when. In addition, we add two more controls to eliminate other, less interesting, factors from the analysis. First, we control for demand signal processing errors by sharing knowledge of the retail demand distribution with all participants. In this sense, our version of game is similar to the “Stationary Beer Game” recently developed by Chen and Samroengraja (1998) for classroom use. Second, we utilize an incentive scheme that avoids problems of collusion and reduced effort over time.

Although this paper is the first to examine the behavioral implications of inventory information sharing, some important theoretical work does exist in this area. Chen (1998) studies the impact of information sharing by quantifying the costs achieved using optimal echelon base-stock and installation base-stock in a multi-echelon inventory system. He found that echelon base-stock policies (which require access to global inventory information) lead to lower cost than installation policies (which require only local information). Since both echelon

and installation base-stock policies imply participants place orders equal to the orders/demand they receive, the bullwhip effect does not emerge in either information state. Other researchers have theoretically examined the costs and benefits of sharing inventory and demand information in two-echelon systems consisting of one manufacturer and one or more retailers (Bourland et al. 1996, Lee et al. 2000, Cachon and Fisher 2000, Gavirneni et al. 1999). Even though the assumed settings and order policies differ between these papers, all conclude that the supplier enjoys larger cost savings from information sharing than his retailers. Our results in Study II confirm that this disparity holds true behaviorally. In particular, when information is available, upstream players enjoy a larger reduction in the variance of orders than downstream players.

3 The Experimental Setting

In this section, we briefly introduce necessary notation, describe the mechanics of the Beer Distribution Game, and explain the protocol followed in conducting the experiments.

3.1 Notation and Experimental Parameters

The Beer Distribution Game mimics the mechanics of a decentralized, periodic review, inventory system with four, serial echelons. Each supply chain team consists of four players, denoted by $i = 1, \dots, 4$, corresponding with the four echelon levels. Figure 1 provides an illustration. Each participant is responsible for placing orders to his upstream supplier and filling orders placed by his downstream customer, over a series of time periods $t = 1, \dots, T$.

-- Insert Figure 1 --

To illustrate the specifics of the game, consider the dynamics of a single team g , where $g = 1, \dots, G$, and G is the number of teams in a particular experiment. Each participant i receives orders placed by its downstream customer ($i-1$) and places orders for additional inventory from

its upstream supplier ($i+1$). For each echelon level and order period, we define inventory as I_t^{ig} , where $I_t^{ig} \geq 0$ represents on-hand inventory and $I_t^{ig} < 0$ represents orders in backlog. Note that when we refer to sharing “inventory” information (e.g., in Study II) we mean sharing I_t^{ig} (i.e., both on-hand inventory and backlog information). We also define the quantity shipped to the downstream customer by S_t^{ig} , the order placed to the upstream supplier by O_t^{ig} , and retail demand in period t by D_t .

Events within each period unfold as follows. First, shipments arrive from upstream suppliers. Second, new orders arrive from downstream customers. In the case of the retailer, the order received at t is simply D_t , retail demand at t . Third, these new orders are filled, if possible, from current inventory and shipped out. If an order cannot be filled, it is placed in backlog. The period ends with each participant placing an order, O_t^{ig} , with his upstream supplier. These are the sole decision variables in the game. Our analysis focuses on how O_t^{ig} varies between roles $i = 1, \dots, 4$ and with the availability of inventory information.

The system is complicated by order processing and product shipment delays between each supplier/customer pair. We assume delay times in keeping with the traditional Beer Distribution Game setup (i.e., two period order and shipment delays between echelons $i = 1, \dots, 3$ and a three period manufacturing delay at echelon $i = 4$; see Figure 1). The order quantity filled and shipped by level i at the end of period t is thus

$$S_t^{ig} = \min\{D_t, \max\{I_{t-1}^{ig} + S_{t-2}^{i+1,g}, 0\}\} \quad \text{for } i = 1 \quad (1)$$

$$= \min\{O_{t-2}^{i-1,g}, \max\{I_{t-1}^{ig} + S_{t-2}^{i+1,g}, 0\}\} \quad \text{for } i = 2, 3 \quad (2)$$

$$= \min\{O_{t-2}^{i-1,g}, \max\{I_{t-1}^{ig} + O_{t-3}^i, 0\}\} \quad \text{for } i = 4 \quad (3)$$

where $I_t^{ig} = I_{t-1}^{ig} + S_{t-2}^{i+1,g} - D_t \quad \text{for } i = 1$

$$= I_{t-1}^{ig} + S_{t-2}^{i+1,g} - O_{t-2}^{i-1,g} \quad \text{for } i = 2, 3$$

$$= I_{t-1}^{ig} + O_{t-3}^{i,g} - O_{t-2}^{i-1,g} \quad \text{for } i = 4$$

Participants are encouraged (through an incentive scheme, described in section 3.2) to choose orders in a manner that minimizes their team's cumulative chain cost. Chain cost for each period t is defined as the sum of holding and backlog costs across the four echelon levels. Using our notation, chain cost through period T for group g , $C^g(T)$, is simply

$$C^g(T) = \sum_{i=1}^4 \sum_{t=1}^T h^i \max\{I_t^{ig}, 0\} - s^i \min\{-I_t^{ig}, 0\} \quad (4)$$

where s^i and h^i are the unit backlog and inventory holding costs incurred at level i , respectively. Note that this supply chain structure controls for three of the four operational causes of the bullwhip effect: inventory allocation (since there are no competing customers and manufacturing capacity is infinite), order batching (since setup times are zero), and price fluctuations (since prices are constant over time).

3.2 Experimental Implementation

The game was implemented in java to run off a web server, with each participant within a team working off a separate computer. Kalidindi (2001) provides a detailed overview of the software. Study participants were drawn from a pool of undergraduate business students at the University of Minnesota currently enrolled in an introductory Operations Management course. Study I and II were both performed on April 30th, 2002.

In both studies, participants arrived in the experimental lab at a pre-determined time and were randomly assigned to a computer terminal, which determined their role i and supply chain assignment g . Participants were told not to communicate with anyone during the experiment, as is standard practice in experimental economics. Once seated, participants were oriented to the

rules and objectives of the game. They were instructed that each role would incur unit holding costs of \$0.50 (h^i) and unit backlog costs of \$1 (s^i) per period, as normally assumed in the board version of the Beer Distribution Game (Sterman 1989a). They were also told that retail demand was uniformly distributed between 0 and 8 cases per period and independently drawn between periods.¹ This allows us to control for the remaining operational cause of the bullwhip effect, demand signal processing.

Each echelon began with an initial inventory level of 12, outstanding orders of 4 for the last two periods, and an incoming shipment of 4 in the next two periods, as shown in Figure 1. Participants were not informed how many periods the experiment would run to avoid end-of-game behavior that might trigger over or under ordering. The actual number of periods was $T = 48$ for all experiments. All experiments also used the same random number seed to generate demand, i.e. D_t , $t = 1, \dots, T$, was identical across groups. This allowed us to isolate variations due to ordering behavior from variations due to different demand streams.

We used a continuous incentive scheme to reward participants for their play, which was also announced at the beginning of the experiment. The incentive scheme provided a base compensation of \$5 for each participant along with a bonus of up to \$20 per participant. The bonus for each member of group g was calculated as follows, based on the cumulative chain profit of the group

$$b^g = \$20 \frac{\max_g(C^g) - C^g}{\max_g(C^g) - \min_g(C^g)} \quad (5)$$

where $C^g = C^g/48$ is defined by (4). This bonus was computed separately for each treatment.

¹One might question whether all subjects understood the meaning of a “uniform distribution.” To overcome this problem, we briefed the subjects about the implications of a uniform (0,8) distribution and wrote the phrase “Demand is U(0,8) meaning (0, 1, 2, 3, 4, 5, 6, 7, or 8) is equally likely in each period” in large letters on the front board. This phrase was left on the board, in clear view, during the entire experiment.

This incentive scheme has a number of attractive properties. First, it provides a continuous incentive for participants to minimize cost in the game. In contrast to previous experimental implementations in which the best-performing group won a fixed prize (e.g. Sterman 1989a), under this payment method, each group has an incentive to control their costs, even if they cannot get into first place by doing so. Second, this design discourages collusion among participants to artificially raise the earnings of all teams together. This is in contrast to the scheme used in Gupta et al. (1998) which paid participants a multiple of their earnings, and thus removed competition between groups. Third, this payment represents the benchmarked cost of an integrated supply chain “firm,” which is a metric of performance often used in industry.

4 Study I: Stationary and Known Demand

In previous experimental studies of the bullwhip effect (e.g., Sterman 1989a, Gupta et al. 1998), participants reacted to demand processes that were both *unknown* and *non-stationary*.² Participants of the traditional Beer Distribution Game (used by Sterman 1989a) often cite the combination of a changing demand pattern and lack of basic distribution information (e.g., demand minimum, maximum, and mode) as the cause of the bullwhip effect. The question examined in our first study is whether the bullwhip effect remains when these operational complications are removed, i.e., when the demand distribution is commonly known and stationary. Subsection 4.1 gives a brief review of existing theory on ordering behavior in this setting, while subsection 4.2 describes our experimental results.

4.1 Theory and Hypotheses

²We define a stationary demand process as one that is drawn from a stable demand distribution, with natural variation but no autocorrelation between draws.

Theoretical research suggests that rational decision-makers will induce the bullwhip effect when the demand distribution is *either* non-stationary *or* unknown to supply chain members. For example, when demand is non-stationary and a simple order-up-to policy is used, a rational decision-maker will adjust order-up-to levels dynamically in each order period. Lee et al. (1997b) show analytically that such a strategy invokes the bullwhip effect in a two-echelon system. This occurs even when the retailer makes order decisions with full knowledge of the underlying (nonstationary) demand distribution. Graves (1999) uncovers a similar result assuming demand follows an autoregressive, integrated moving average (ARIMA) process. In this case, he develops an explicit expression for the magnitude of amplification, which appears independent of the level of information provided to upstream stages.

When the demand distribution is stationary but unknown, it makes sense to use a stationary forecasting technique to estimate demand. Chen et al. (1998) show that the bullwhip effect can occur when managers follow a simple order-up-to policy that is updated in each period based on current demand information. Their results suggest that the bullwhip effect will be largest during transient periods, when little demand information is available, and may diminish once enough demand information is available to provide an accurate forecast. Anderson and Fine (1998) give further credence to the hypothesis that demand signal updating induces the bullwhip effect through an analysis of real industry data.

If the demand distribution is both stationary and commonly known, theory suggests that no bullwhip effect will exist. Chen (1999) shows that a base-stock policy is optimal in a serial inventory system with fixed order and shipment delays. This policy was also shown to be optimal for serial systems without delays by Clark and Scarf (1960) and Federgruen and Zipkin (1984). A base-stock policy implies that participants place orders equal to the orders they

receive. Thus we hypothesize that rational decision makers will not induce the bullwhip effect in our first study where demand is known and stationary.

Hypothesis 1: *The bullwhip effect will not occur when the demand distribution is known and stationary.*

As we see in the next subsection, our experimental results are not consistent with this prediction.

4.2 Experimental Results

Forty-four subjects participated in this study, divided across eleven supply chain teams. Figure 2 illustrates an example of the trend of orders placed for a typical team.

-- Insert Figure 2 --

Recall that the bullwhip effect is defined as an increase in the variance of orders placed at a given echelon level, relative to the orders received at that level. This implies that the variance of orders is amplified as one moves up the supply chain. Hypothesis 1 implies instead that the variance of orders across levels will remain stable. Figure 3 graphs the variance of orders placed for the forty-four participants. By inspection, it appears that some amplification in order variance occurs as one moves up the chain. If we look at the ratio of average variances between roles, we find that the wholesaler/distributor pairing exhibit the largest magnitude of amplification: $\frac{\bar{\sigma}_2^2}{\bar{\sigma}_1^2} = 1.73$, $\frac{\bar{\sigma}_3^2}{\bar{\sigma}_2^2} = 2.11$, $\frac{\bar{\sigma}_4^2}{\bar{\sigma}_3^2} = 1.48$, where $\bar{\sigma}_i^2$ is the average variance for role i . Rachel: I keeping the next two sentences as is. The experimental results of Sterman (1989a) show a similar pattern (his ratios, calculated from Table 3 of Sterman (1989a), are $\frac{\bar{\sigma}_2^2}{\bar{\sigma}_1^2} = 1.77$, $\frac{\bar{\sigma}_3^2}{\bar{\sigma}_2^2} = 1.96$, $\frac{\bar{\sigma}_4^2}{\bar{\sigma}_3^2} = 1.60$).

-- Insert Figure 3 --

4.2.1 Evidence of the Bullwhip Effect

A simple nonparametric sign test confirms that amplification is present in this setting. The test works as follows. For each supply chain, we code an increase in the variance of orders placed between each role as a success, and a decrease as a failure. More details on the sign test can be found in Seigel (1965), p. 68 and Table D. The data reveals an 82% success rate; 82% of the observations involve an increase in variance of orders placed between roles. This is significantly different than the null hypothesis (50%) implied by Hypothesis 1 ($N=33$, $x=6$, $p<0.0001$). This is a conservative estimate since, if the variance at the wholesale level is higher than variance at the retail level (for example) then the probability that variance at the distributor level is greater than variance at the wholesale level is actually *less than* 50% due to regression to the mean (although the exact rate is difficult to estimate).

This result leads us to reject Hypothesis 1. The bullwhip effect still exists when retail demand is stationary and commonly known. However, it is unclear why decision-makers act in this fashion. In the next subsection, we test the existence of one type of decision bias identified by Sterman (1989a) that may contribute to this seemingly irrational behavior.

4.2.2 Evidence of Under-Weighting the Supply Line

Sterman (1989a) found that participants fail to take proper assessment of their current outstanding order level when making order decisions subject to an unknown and non-stationary retail demand distribution. In other words, they under-weight the supply line. Our purpose here is to test whether this decision bias holds when the retail demand distribution is known and stationary. One might expect participants to pay closer attention to the supply line in this more stable setting. Therefore, we seek to test the following hypothesis.

Hypothesis 2: *Participants will not under-weight the supply line when demand is known and stationary.*

Following Sterman's analysis, we test participants' perception of the supply line as a whole (i.e., all outstanding orders) by running a set of regressions, one for each participant in our experiment. Each regression compares orders placed in period t against the participant's on-hand inventory level, orders received and outstanding orders in period t . The regression equation for a given participant i , in group g , at time t , is

$$O_t^{ig} = \max \{0, \alpha_0 + \alpha_I I_{t-1}^{ig} + \alpha_R R_t^{ig} + \alpha_S S_t^{ig} + \alpha_N N_t^{ig} + \alpha_t t + \varepsilon\} \quad (6)$$

where I_{t-1}^{ig} denotes the inventory held last period, R_t^{ig} denotes orders received from the downstream customer (i.e., $R_t^{ig} = O_{t-2}^{i-1,g}$), S_t^{ig} describes the shipments received from his upstream supplier (defined by equations 1-3), N_t^{ig} denotes the participant's total outstanding orders (i.e., $N_t^{ig} = O_{t-1}^{ig} - \min\{0, I_t^{i+1,g}\} + S_t^{i+1,g} + S_{t-1}^{i+1,g}$), and t serves as a control for any time trends.³

If the bullwhip effect did not exist, a participant's order quantity would simply equal the orders received from his downstream customer. In our regression equation, this implies $\alpha_R = 1$ after the initial transient period ends (i.e., once the participant has a chance to either deplete excess initial inventory or raise the initial inventory level to the order-up-to level). An exogenous shock increasing inventory on-hand in the supply line, or in shipments received, should reduce orders by one unit in order to retain the order-up-to-level. This implies $\alpha_I = \alpha_N = \alpha_S = -1$. Note that if the participants were ordering according to the optimal order-up-to policy, α_0 would reflect the optimal order-up-to level.

³The astute reader will notice that the regressions we run to test Sterman's hypothesis are simpler than the ones used in his paper. Sterman's original analysis included a number of assumptions concerning the heuristics used by his participants. For example, he assumed that participants anchored in their choice of desired stock, with the anchor estimated separately for each subject. Thus each subject was thought to order up to a fixed inventory level in each time period. Sterman also assumed participants used adaptive expectations to forecast the orders they would receive. Neither of these assumptions seemed reasonable in our setting, where participants knew the demand distribution facing the retailer. Our regression expression (6) is an attempt to capture Sterman's insight without relying on his assumptions of anchoring and forecasting.

If participants are accurately accounting for the supply line, the coefficient on orders outstanding should be the same as the coefficient on inventory. If Sterman's conjecture holds (i.e., participants are under-weighting the supply line), then we should find $\alpha_N > \alpha_I$.

We ran this regression for each of the 44 participants in our study (4 roles x 11 groups), using the data from all but the last two periods of the game.⁴ The average R^2 (adjusted) statistic over all regressions was 0.68. Overall, the regressions were highly significant with only one participant's F-statistic indicating his choices could be equally well explained by a constant order amount (and that participant's regression was marginally significant, with $p=.08$). The coefficient for on-hand inventory (α_I) was significantly higher than the optimal value of -1 for all but 8 of the 44 regressions (at the $p<.05$ level or better) with an average value of $-.236$. Similarly, the coefficient on the orders received term, α_R , was significantly lower than the optimal value of $+1$ for all but 10 of the 44 regressions (again at the $p<.05$ level or better) with an average value of $.331$.

With respect to underweighting the supply line, we found that the average value for α_N was -0.029 compared to -0.236 for α_I . All of the 44 participants placed less weight on their supply line than on their inventory positions ($\alpha_N > \alpha_I$), and thus under-weighted their supply line. We can easily reject Hypothesis 2 (which implies an equal likelihood of $\alpha_N > \alpha_I$ and $\alpha_N < \alpha_I$) using a sign test ($N=44$, $x=0$, $p<0.0001$). Thus, we conclude that a supply line underweighting bias still occurs when retail demand is stationary and known. It is interesting to note that our estimated regression parameters are similar to those derived by Sterman (1989a). For example, Sterman reports an average inventory coefficient of -0.26 (versus our -0.236) and an

⁴In the last two periods of the game, the shipments that an upstream counterpart could send were not determined since they depended, in part, on the orders that counterpart would have placed had the game continued.

average supply line coefficient of -0.0884 (versus our -0.029). It appears that having a stable, and known, demand distribution does little to alleviate subjects' tendency to underweight the supply line.

Previous experimental research offers some clues to why participants continue to underweight the supply line even when demand is known and stationary. In a literature review, Sterman (1994) notes that dynamic settings render decision-making difficult, even when only one decision-maker is involved, due to reduced saliency of feedback. Saliency refers to the strength of the tie between feedback and the decision (Hogarth 1987). Diehl and Sterman (1995), Paich and Sterman (1993), Sterman (1989b) and Sterman (1989c) provide specific examples and experimental evidence in individual decision-making tasks. When decisions are made in a decentralized fashion across multiple parties, such as in our supply chain setting, the interaction of decisions and outcomes further degrades the saliency of feedback (Hogarth 1987). Other authors suggest that the type of feedback available is critical to the learning process, and that outcome feedback (rather than process feedback) causes inefficiencies (e.g., Sengupta and Abdul-Hamid 1993).

One type of feedback missing in this and many other supply chain settings is a forewarning of when upstream suppliers (or downstream customers) are running short on inventory. Without this information, a participant cannot be certain that the shipments they receive will correspond to orders they actually placed. A decision-maker's connection between cause and effect (here, orders requested and shipments received) is broken whenever his supplier is out of stock. Having access to the inventory levels of upstream members could improve a decision-maker's ability to anticipate supply shortages and thus increase the salience of feedback. If the participant knew when (and why) stock-outs occurred, his understanding of the relationship between orders placed and orders received might improve. This information might

also combat the supply line under-weighting tendency just observed, since it provides a clear view of one critical component of a participant's supply line (his supplier's inventory level). This is the behavioral motivation behind our second study. The second study is of interest for operational reasons as well, as it sheds light on the behavioral impact of inventory tracking systems being implemented in practice.

5 Study II: Sharing Dynamic Inventory Information

The purpose of our second study is to measure what improvement is gained by sharing current inventory information across the supply chain. Conditions in this experiment were identical to Study I except that the participants also had access to dynamic inventory information. This information was displayed in a bar chart with four bars representing the inventory positions of the four members of their group. The chart was automatically updated at the beginning of each period so all members had access to the same information when making their order decisions. Before discussing the results of the experiment, we briefly outline our hypotheses.

5.1 Hypotheses

Our main hypothesis is that sharing inventory information will help reduce the bullwhip effect. We can measure reductions to the bullwhip effect in two ways: reduced order oscillations at a given supply level, and reduced amplification across levels. The first measure reflects efficiency improvements to the chain as a whole, while the second focuses on the rate of change between levels. Theoretical research suggests that inventory information can improve the procurement process for a supplier serving one or more retailers (e.g., Bourland et al. 1996, Lee et al. 2000, Cachon and Fisher 2000, Gavirneni et al. 1999) and for multiple decision makers in a serial supply chain (Chen 1998). We hypothesize that inventory information will

help in our setting as well by reducing order oscillations throughout the chain. This leads to the following hypothesis.

Hypothesis 3: *Sharing dynamic inventory information across the supply chain will decrease the level of order oscillation.*

Whether the amplification of order oscillations will decrease between levels as a result of inventory exposure is less clear. Theory suggests the bullwhip will not occur when the demand distribution is known and stable, whether or not inventory information is shared (Chen 1998). However, we saw in Study I that the bullwhip effect does appear when inventory information is not available. We hypothesize that inventory information may help alleviate the supply line under-weighting bias reported in section 4.2.2 by increasing the saliency of feedback between orders placed and shipments received and thus decreasing the amplification of orders. This leads to the following two hypotheses.

Hypothesis 4: *Sharing dynamic inventory information across the supply chain will decrease the amplification of order oscillation between each supply chain level.*

Hypothesis 5: *Sharing dynamic inventory information will cause participants to no longer under-weight the supply line.*

Another interesting question to ask is who benefits the most from information sharing.⁵ Theoretical research of two-echelon inventory systems (e.g., Bourland et al. 1996, Lee et al. 2000, Cachon and Fisher 2000, Gavirneni et al. 1999) suggests that manufacturers enjoy most of the operational benefits. Indeed, once inventory information is shown, the improvement in demand information is most profound for upstream participants. Access to downstream information allows suppliers to manage their orders based on echelon inventory level rather the

⁵Of course, the fact that one member yields greater operational improvements (such as reduced order oscillations) does not guarantee this same member will enjoy a larger share of the monetary gains. In a market economy, these cost improvements could very well be transferred to a more powerful channel member or passed on to final customers in the form of lower prices.

order quantity placed by his immediate customer (Chen 1998), which effectively eliminates the order amplification component of the bullwhip effect.

On the other hand, for downstream participants, inventory information offers an opportunity to better anticipate supply shortages and possibly increase their understanding of the ordering process used by their suppliers. This may also help correct a retailer's tendency to under-weight the supply line. In particular, it may reduce his tendency to over-order when shipments begin to fall short of previous order levels. If this is the main benefit of inventory information, we may find downstream partners have more to gain.

To test the relative benefits achieved by upstream and downstream participants, we divide the supply chain into two segments. This leads to the following, competing, hypotheses.

Hypothesis 6a: *Sharing dynamic inventory information across the supply chain will lead to a greater reduction in order oscillations for manufacturers and distributors than for retailers and wholesalers.*

Hypothesis 6b: *Sharing dynamic inventory information across the supply chain will lead to a lower reduction in order oscillations for manufacturers and distributors than for retailers and wholesalers.*

The next subsection tests these hypotheses by comparing a new set of experiments against the results of Study I.

5.2 Experimental Results

Forty-four different subjects, divided across eleven supply chain teams, took part in this second study. Figure 4 illustrates the variance of orders placed for the participants. It appears that the bullwhip effect still exists, but may be less severe than that observed in Study I (i.e., Figure 3).

--Insert Figure 4--

5.2.1 Impact on Overall Chain

Focusing first on Hypothesis 3, we use a nonparametric 2-tailed Mann-Whitney U test (also called the Wilcoxon test)⁶ to compare how the oscillation component of the bullwhip effect compares across the two treatments. The test confirms that order variances are significantly less when inventory is known, compared with the control treatment ($n=44$, $m=44$, $z=1.92$, $p=0.028$), providing support for the hypothesis.

While order variances may be lower, it is interesting to note that the bullwhip effect still persists when inventory information is shown. Using the same sign test discussed in section 4.2.1, we find that the variability of orders placed between each role increased 69% of the time (i.e., exhibited a 69% success rate) when inventory information was commonly known. This is significantly different than the 50% success rate of the null hypothesis if no amplification existed ($N=33$, $x=10$, $p=0.0107$), but is lower than the 80% rate of increase observed previously. This suggests that the amplification component of the bullwhip effect is reduced, but still present, when inventory information is commonly known.

To examine this amplification aspect of the bullwhip effect more rigorously, we perform a Mann-Whitney U test on the ratio of the variances between each of the two roles (e.g. the variance of a wholesaler's orders divided by the variance of his retailer's orders) between the two treatments. A Mann-Whitney U test using all the levels of the chain (33 observations from the control treatment and 33 observations from the inventory-shown treatment) suggests no significant difference ($n=33$, $m=33$, $z=0.468$, $p=0.320$) between the two treatments. However, a similar test performed separately on each customer/supplier link shows some evidence of amplification reduction between roles. In particular, there is a significant reduction in amplification between the wholesaler and distributor ($n=11$, $m=11$, $z=1.609$, $p=0.044$) across the two treatments. In contrast, the retailer/wholesaler ($n=11$, $m=11$, $z=1.64$, $p=0.435$) and

⁶Seigel (1965) discusses the Mann Whitney U test starting on page 116.

distributor/manufacturer ($n=11$, $m=11$, $z=1.145$, $p=0.125$) links show no significant difference between treatments. These tests suggest that while the impact of information sharing on reducing amplification is not significant overall (rejecting Hypothesis 4), it may offer some benefit to upstream players. We revisit this issue in section 5.2.3.

5.2.2 Impact on Supply Line Weighting

To test whether participants continue to under-weight the supply line in this setting, we ran the same individual regression on each of the 44 participants in this study. The average R^2 (adjusted) statistic for these regressions was 0.561. The average inventory weight was -0.193 , while the average weight placed on the supply line was -0.029 . Forty-two out of 44 participants under-weighted the supply line. As before, this pattern of results is significantly different than would be expected if the supply line were being weighted equally as inventory using a sign test ($N=44$, $x=2$, $p<.0001$). These results are similar to those in the first study, allowing us to reject Hypothesis 5. Our results suggest that participants' under-weighting of the supply line is robust to the additional information in this experiment (i.e., the inventory position of other firms). This result supports the work of Diehl and Sterman (1995) and Sterman (1998b), who found that additional information does not counteract the under-weighting tendency in other experimental settings. It appears that downstream customers do not use inventory information to their full advantage.

If supply line under-weighting is still prevalent when inventory information is shown, then what is causing the improvement in performance? The next subsection suggests that while inventory information is not being used by downstream members to adjust for outstanding orders, it could be used by upstream members of the supply chain to anticipate and adjust for

downstream members' orders. This use counteracts the under-weighting of the supply line and improves overall performance.

5.2.3 Impact on Upstream and Downstream Members

Table 1 compares the average order variance for the two treatments by role. Here we see that the magnitude of improvement appears much larger for upstream supply chain members (i.e., the distributor and manufacturer). To test the significance of this improvement for upstream versus downstream members, we once again use a set of Mann-Whitney U tests. Grouping distributors and manufacturers together reveals that inventory information leads to a significant reduction in the variance of orders at upstream sites ($n=22$, $m=22$, $z=1.82$, $p=.043$). Grouping retailers and wholesalers together, reveals an insignificant difference between the treatments ($n=22$, $m=22$, $z=1.24$, $p=.110$).

-- Table 1 here --

Table 1 also reports the percentage reduction in order variance enjoyed by each role with the introduction of inventory information, both compared with the first study and compared with the benchmark of no bullwhip effect. These calculations highlight the asymmetric improvement observed between upstream and downstream members of the supply chain.

These results suggest that members near the beginning of the chain exhibit a different impact from inventory information than those near the end. In particular, our results suggest that inventory information may be more useful as one moves further away from end-user demand. This behavior is consistent with Hypothesis 6a. These results on the behavioral benefits of inventory information on upstream versus downstream members concur with previous theory (e.g., Bourland et al. 1996, Lee et al. 2000, Cachon and Fisher 2000, Gavirneni et al. 1999).

Upstream members exhibit a significant reduction in order oscillations, while downstream members show relatively little improvement.

Note that one key difference between competing Hypotheses 6a and 6b is the extent to which they predict under-weighting of the supply line. Hypothesis 6b suggests that sharing inventory information will trigger retailers and wholesalers to stop under-weighting the supply line due to upstream information. We saw in section 5.2.2 that this was not the case. Our individual results here suggest that inventory information is of limited value in alleviating this behavioral cause of the bullwhip effect. Nonetheless, inventory information counteracts this bias and improves performance by allowing manufacturers and distributors to anticipate and interpret orders placed by their downstream customers.

6 Discussion and Conclusion

This paper reports the results of two experimental studies on the behavioral causes of the bullwhip effect. We find participants continue to exhibit the bullwhip effect (the amplification of oscillation of orders higher in the supply chain) even under conditions where it rationally should not occur. This suggests that cognitive limitations contribute to the bullwhip effect, even in ideal and controlled settings like the lab. We also observe that transmitting dynamic inventory information lessens the bullwhip effect, particularly at higher echelon levels. We argue that this information allows upstream members to better interpret orders on the part of their customers and prevents them from overreacting to fluctuations when placing their own orders. These studies provide several important implications for managerial practice and for guiding the development of inventory theory.⁷

⁷Some might claim that the participants in our experiment were from a different subject pool than people who actually make inventory decisions for a living. This is a problem common to most experimental work. Our

Results from our first study suggest that the bullwhip effect is not solely a result of operational complications such as seasonality or unpredictable demand trends. Indeed, the effect remains even under the most optimistic demand scenario, when the demand distribution is stationary and commonly understood. The reason has to do with cognitive limitations on the part of managers and difficulties inherent in managing a complex dynamic system. One particular limitation appears to be that participants under-weight the supply line, tending to discount orders they have placed but which have not been delivered.

Many companies are in the process of implementing integrated information systems to manage their distribution processes. Our results indicate that these systems have the potential to decrease the severity of the bullwhip effect. Further analysis revealed this potential is significantly larger for upstream members. This result suggests several implementation philosophies, which warrant further study. For example, these results lead us to speculate that the critical part of an inventory-sharing information system is not communicating the inventory position of the manufacturer to the retailer, but rather communicating the inventory position of the retailer to the manufacturer. This implies that the biggest “bang for the buck” from these systems may lie in tracking and sharing downstream inventory information, i.e. information closest to the final customer. Since the cost of inventory tracking is quite high, particularly at manufacturing sites, one might do well to implement such systems first at the retail and wholesale level. We predict diminishing returns as such systems are implemented further up the

response is that today’s business students are tomorrow’s inventory professionals. Presumably experiments run with students pull from the same population sample as inventory professionals, and thus we are not selecting a different set of people to study. Of course, there is a question about training. Since inventory professionals have had more practice and experience in this problem, perhaps they would not exhibit the same biases as (untrained) students. To test this impact we ran the first treatment (study I) with members of APICS (American Production and Inventory Control society, a professional organization of operations managers). Outcomes from those experiments showed the same patterns as those reported here.

supply chain. We also conjecture that an inventory-sharing system may still be effective even if upstream firms are reluctant or unable to share their inventory positions with downstream firms.

With respect to inventory theory, our results point to the need for more theoretical research that incorporates the biases of individual decision-makers. In our experience, these factors are often overlooked when assessing the expected benefits of investments in information technology. Our results suggest that models based on unboundedly rational actors may not reflect actual decision-making behavior and thus may undervalue the benefit from investments in information technology. We hope this research stimulates new theoretical models which better capture supply line decision biases, the boundedly rational nature of supply chain managers, and the impacts of these limitations not only on actions taken within one particular informational structure, but the choice of informational structure itself.

References

- Anderson, E. and C. Fine (1998), "Business Cycles and Productivity in Capital Equipment Supply Chains," chapter 13 in *Quantitative Methods for Supply Chain Management*, S. Tayur, Ganeshan, and Magazine (eds.), Kluwer Academic Publishers.
- Anderson, E. and D. Morrice (2000) "A Simulation Game for Teaching Service-Oriented Supply Chain Management: Does Information Sharing Help Managers with Service Capacity Decisions?" *Production and Operations Management*, **9**, 1, 40-55.
- Bourland, K., S. Powell, and D. Pyke, (1996) "Exploiting Timely Demand Information to Reduce Inventories," *European Journal of Operational Research*, **92**, 239-253.
- Cachon, G., (1999) "Managing Supply Chain Variability with Scheduled Ordering Policies," *Management Science*, **45**, 843-856.
- Cachon, G. and M. Fisher, (2000) "Supply Chain Inventory Management and the Value of Shared Information," *Management Science*, **46**, 1032-1048.
- Cachon, G., and M. Lariviere, (1999) "Capacity Choice and Allocation: Strategic Behavior and Supply Chain Performance," *Management Science*, **45**, 1091-1108.
- Chen, F., (1998) "Echelon Reorder Points, Installation Reorder Points, and the Value of Centralized Demand Information," *Management Science*, **44**, S221-S234.
- Chen, F., (1999) "Decentralized Supply Chains Subject to Information Delays," *Management Science*, **45**, 1016-1090.
- Chen, F., Z. Drezner, J.K. Ryan, and D. Simchi-Levi, (1998) "The Bullwhip Effect: Managerial Insights on the Impact of Forecasting and Information on Variability in a Supply Chain," chapter 14 in *Quantitative Models for Supply Chain Management*, Tayur, Ganeshan, and Magazine (eds.), Kluwer Academic Publishers.
- Chen, F. and R. Samroengraja, (2000) "The Stationary Beer Game," *Production and Operations Management*, **9**, 19-30.
- Clark, A.J. and H.E. Scarf, (1960) "Optimal Policies for a Multi-Echelon Inventory Problem," *Management Science*, **6**, 475-490.
- Diehl, E. and J.D. Sterman, (1995) "Effects of Feedback Complexity in Dynamic Decision Making," *Organizational Behavior and Human Decision Processes*, **62**, 198-215.
- Federgruen, A. and P. Zipkin, (1984) "Computational Issues in an Infinite Horizon, Multi-Echelon Inventory Model," *Operations Research*, **32**, 818-836.
- Forrester, J., (1958) "Industrial Dynamics: A Major Breakthrough for Decision Makers," *Harvard Business Review*, **36**, 37-66.

Gavirneni, S., R. Kapuscinski, and S. Tayur, (1999) "Value of Information in Capacitated Supply Chains," *Management Science*, **45**, 16-24.

Graves, S.C. (1999) "A Single-Item Inventory Model for a Nonstationary Demand Process," *Manufacturing and Service Operations Management*, **1**, 50-61.

Gupta, S., J. Steckel and A. Banerji, (1998) "Dynamic Decision Making in Marketing Channels: An Experimental Study of Cycle Time, Shared Information and Consumer Demand Patterns." Stern School of Business working paper, New York University. **RACHEL – UPDATE CITE.**

Hogarth, R., (1987). *Judgment and Choice*, 2nd edition, Wiley, New York.

Kahneman, D. and D. Lovallo, (1993) "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking," *Management Science*, **39**, 17-31.

Kalidindi, V.S. (2001) "The Web Based Beer Distribution Game: An Application", unpublished Masters Thesis, Department of Mechanical Engineering, University of Minnesota.

Kaminsky, P. and D. Simchi-Levi, (1998) "A New Computerized Beer Distribution Game: Teaching the Value of Integrated Supply Chain Management," in *Global Supply Chain and Technology Management*, H. Lee and S.M. Ng (editors), The Production and Operations Management Society series in Technology and Operations Management, **1**, 216-225.

Lee, H., P. Padmanabhan and S. Whang, (1997a) "Information Distortion in Supply Chains," *Sloan Management Review*, **38**, 93-102.

Lee, H., P. Padmanabhan and S. Whang, (1997b) "Information Distortion in a Supply Chain: The Bullwhip Effect," *Management Science*, **43**, 546-558.

Lee, H., K. So and C. Tang, (2000) "The Value of Information Sharing in a Two-Level Supply Chain," *Management Science*, **46**, 626-643.

Paich, M. and J.D. Sterman, (1993) "Boom, Bust, and Failures to Learn in Experimental Markets," *Management Science*, **39**, 1439-1457.

Scheck, S., (1998) "Net Tools Could Save Automakers \$1 Billion," *Electronic News*, September 14, 104.

Schotter, A. and K. Weigelt, (1992) "Behavioral Consequences of Corporate Incentives and Long-Term Bonuses: An Experimental Study," *Management Science*, **38**, 1280-1298.

Schweitzer, M. and G. Cachon, (2000) "Decision Biases and Learning in the Newsvendor Problem: Experimental Evidence," *Management Science*, **46**, 404-420.

Seigel, S. (1965) *Nonparametric Statistics for the Behavioral Sciences*, Wiley, New York.

Sengupta, K. and T. Abdel-Hamid, (1993) "Alternative Conceptions of Feedback in Dynamic Decision Environments: An Experimental Investigation," *Management Science*, **39**, 411-428.

Sogomonian, A. and C. Tang, (1993) "A Modeling Framework for Coordinating Promotion and Production Decisions within a Firm," *Management Science*, **39**, 191-203.

Stein, T., (1998) "SAP Targets Apparel," *Information Week*, 140, May 4, 1998.

Sterman, J.D., (1989a) "Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment," *Management Science*, **35**, 321-339.

Sterman, J.D., (1989b) "Misperceptions of Feedback in Dynamic Decision Making," *Organizational Behavior and Human Decision Processes*, **43**, 301-335.

Sterman, J.D., (1989c) "Deterministic Chaos in an Experimental Economic System," *Journal of Economic Behavior and Organization*, **12**, 1-28.

Sterman, J.D., (1992) "Teaching Takes Off: Flight Simulators for Management Education," *OR/MS Today*, **19**, (5), 40-44.

Sterman, J.D., (1994) "Learning in and About Complex Systems," *System Dynamics Review*, **10**, 291-330.

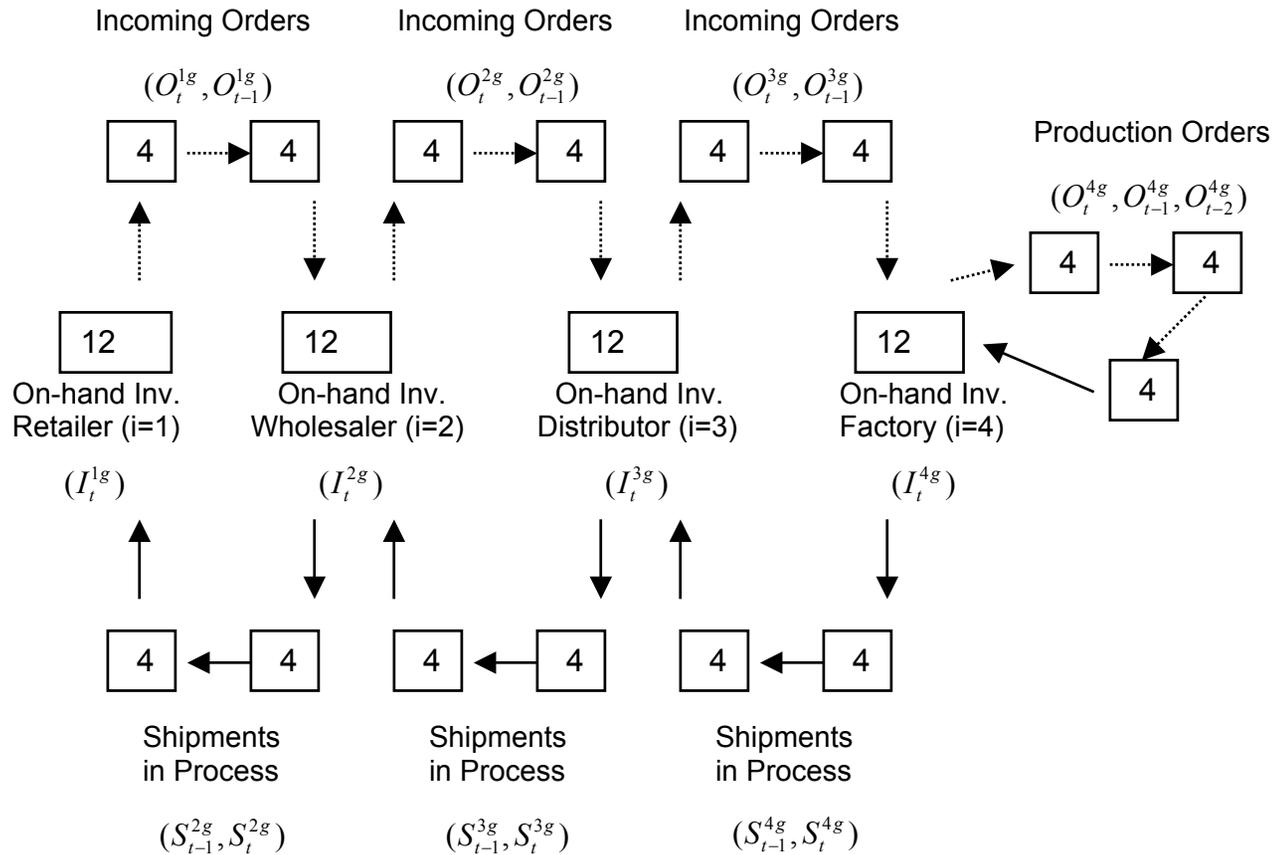


Figure 1: Distribution System Used in the Beer Distribution Game

Note: The numbers listed indicate the initial quantities (of outstanding orders, on-hand inventory, and shipments in process) at the start of the game.

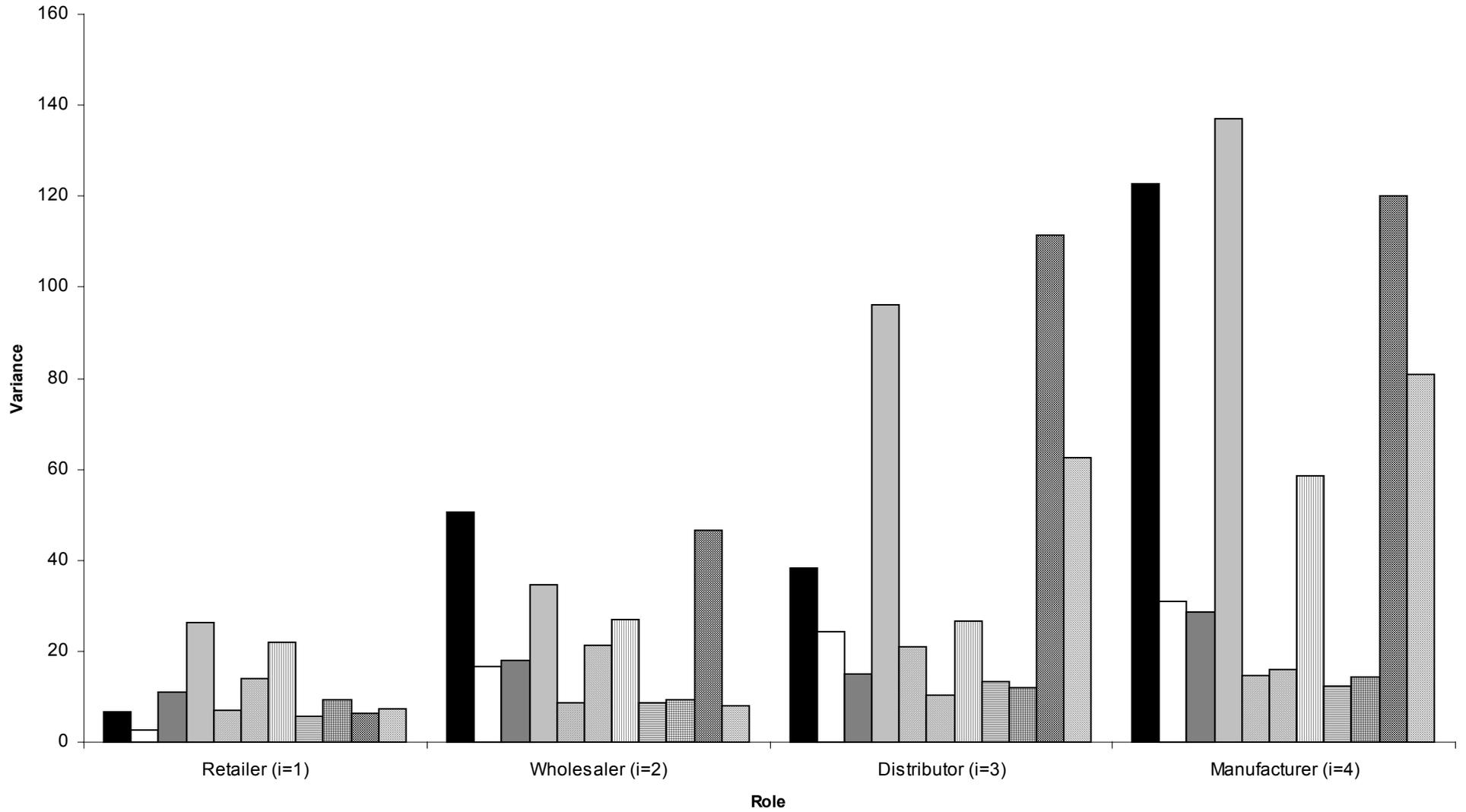


Figure 3: Variance of Orders for Study I

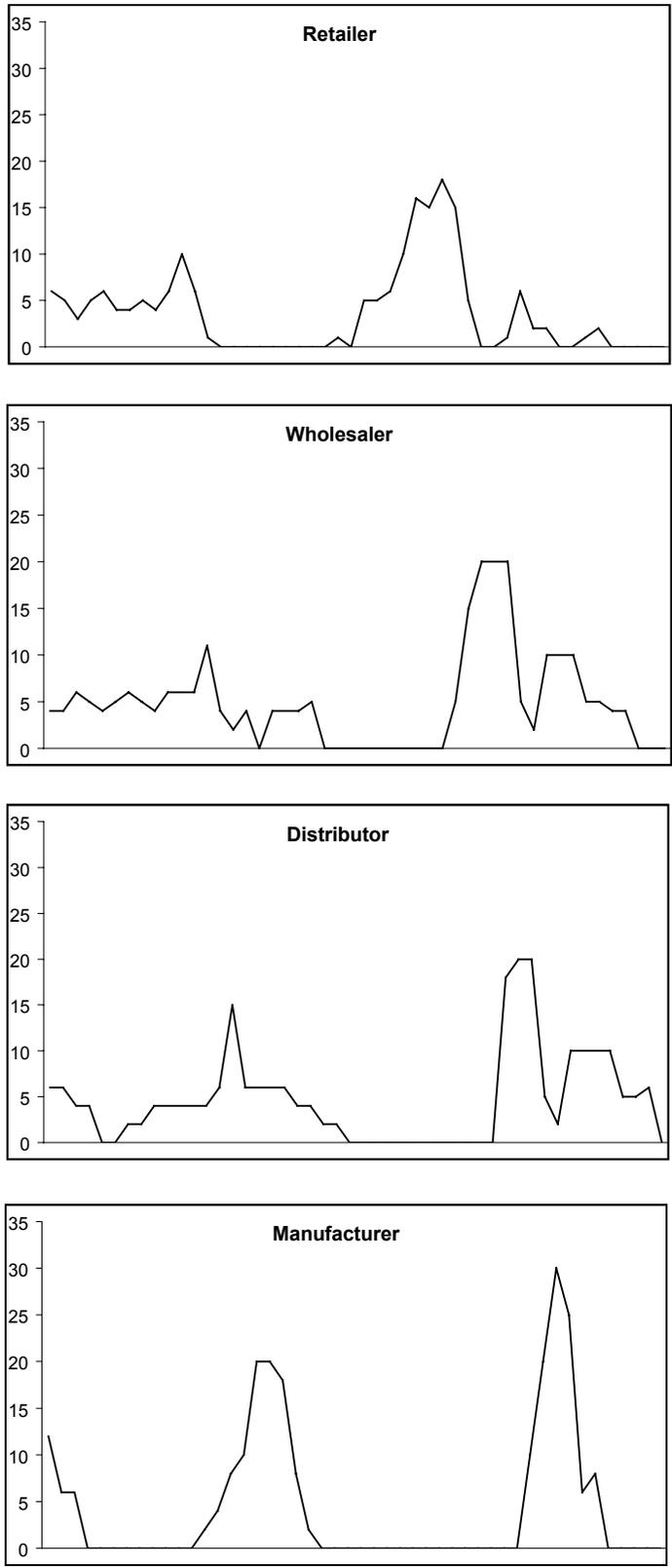


Figure 2: Orders Placed for Group 1 in Study I (Base-Case)

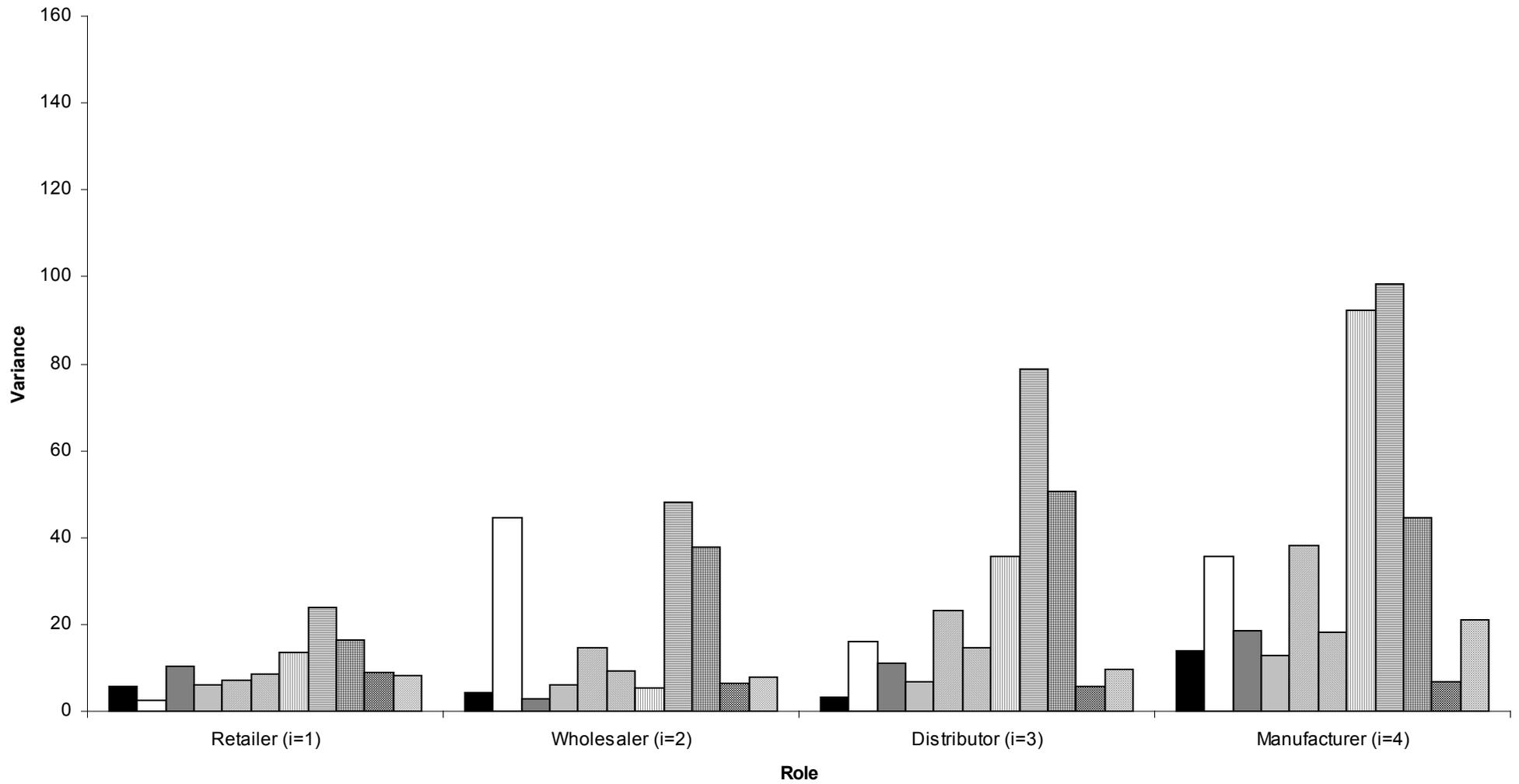


Figure 4: Variance of Orders for Study II (Inventory Shown)

Role	Average variance of orders (σ_i)		Improvement (%)	Improvement relative to 'no bullwhip' benchmark*
	Base case (Study I)	Inventory Shown (Study II)		
1. Retailer	10.73	10.13	5.6%	11.1%
2. Wholesaler	18.56	16.98	8.5%	11.9%
3. Distributor	47.83	23.18	51.5%	58.0%
4. Manufacturer	57.95	36.37	37.2%	41.0%

*In the 'no bullwhip' benchmark, the order variance at each role matches the variance of customer demand (which is 5.33 for our uniform (0,8) distribution). This improvement statistic is thus: $(\sigma_{i,1} - \sigma_{i,2}) / (\sigma_{i,1} - 5.33)$, where $\sigma_{i,k}$ is the average variance for role i in study k.

Table 1: Comparison of average order variance, by role, across the two studies.