The impact of ordering behavior on order-quantity variability: A study of forward and reverse bullwhip effects

Ying Rong • Zuo-Jun Max Shen • Lawrence V. Snyder

Department of Industrial and Systems Engineering, Lehigh University, 200 West Packer Avenue, Mohler Lab, Bethlehem, PA 18015, USA
Department of Industrial Engineering and Operations Research, University of California–Berkeley, 4121 Etcheverry Hall, Berkeley, CA 94720, USA
Department of Industrial and Systems Engineering, Lehigh University, 200 West Packer Avenue, Mohler Lab, Bethlehem, PA 18015, USA
ying.rong@lehigh.edu • shen@ieor.berkeley.edu • larry.snyder@lehigh.edu

We use simulation and live beer game experiments to test the effect of various ordering behaviors under a range of settings for the supply process in a serial supply chain. We find that the bullwhip effect (BWE) does not occur in all cases; moreover, we find evidence of a reverse bullwhip effect (RBWE), in which order variability increases as one moves downstream in the supply chain. We find that the pattern of standard deviations across the supply chain depends heavily on the relative weight that players place on their on-hand and supply-line inventory when choosing an order quantity. In general, demand uncertainty and an underweighting of the supply line cause the BWE, while safety stock inventory dampens it. In contrast, supply uncertainty and an overweighting of the supply line, as well as an over-reaction to random capacity shocks, cause the RBWE.

1. Introduction

The bullwhip effect (BWE) describes the amplification of order variability as one moves upstream in the supply chain. The BWE was formally introduced and analyzed by Lee et al. [18], and has since drawn extensive attention from both academia and industry. However, recent evidence suggests that the BWE does not prevail in general. Baganha and Cohen [2] study the quantity of shipments from manufacturers and that of sales from wholesalers and retailers in the USA from 1978 to 1985. They conclude, by examining the coefficient of variation, that it is the wholesalers rather than the manufacturers who see the largest variance of demand (i.e., sales to retailers). This implies that the wholesalers actually smooth the orders received from the retailers rather than amplify them. Moreover, Cachon el at. [5]
perform a detailed empirical study at the industry level and show that only 47% of industries studied display the BWE, while the remaining 53% show the reverse.

Similarly, although a number of studies have confirmed the presence of the BWE in an experimental setting using the well known beer game [7, 10, 11, 12, 13, 15, 21, 25, 27], some of these studies [11, 12, 15, 27] find a portion of trials in which the opposite effect occurs.

In this paper, we use beer game and simulation experiments to test the ubiquity and magnitude of the BWE when the supply capacity is constrained. When supply is uncertain, we find evidence of a reverse bullwhip effect (RBWE), in which order variability increases as one moves downstream in the supply chain. As a general claim, we can say that demand uncertainty tends to cause the BWE, while supply uncertainty tends to cause the RBWE. However, both of these tendencies are subtly influenced by players’ ordering behaviors, which we explore in much greater detail below.

Consumer purchasing behavior following hurricanes Katrina and Rita in 2005 provides anecdotal evidence of the RBWE. Information Resources Inc. (IRI) [1] provides sales data for consumer packaged goods (CPG) across the United States between the landfalls of the two hurricanes and finds a significant fluctuation in demand for certain products, such as food and water, during this time. This suggests that demand volatility is greater than supply volatility—the reverse of the BWE. A Wall Street Journal article from the time [14] describes similar patterns in gasoline buying:

Typically, car tanks are about one-quarter full. If buyers start keeping car tanks three-quarters full, the added demand would quickly drain the entire system of gas supplies. Dan Pickering, president of Houston-based Pickering Energy Partners, says: “What you don’t have in the system is the ability to run every car full of gas. If you get a hoarding mentality among the consumers, then it tightens the system even further. Fear of shortage begets the shortage. It becomes a vicious circle.”

Indeed, long gas lines were reported throughout the U.S. following Katrina [20]. In this case, customers reacted to the sudden drop in capacity, skewing their demand pattern. This suggests that the RBWE, rather than the BWE, may occur during times of reduced capacity.

In order to simulate this situation, we use a variant of the beer game to test the responses people make in a serial supply chain subject to random capacity shocks, i.e., decreases in capacity. (Previous beer game experiments have used uncertainty at the demand side
One recurring question with any beer game experiment is whether the participants are sufficiently representative of ordering managers in actual supply chains. First, the majority of participants in published beer game studies are either undergraduates or MBA students, and many were not familiar with the BWE or the beer game before participating in the study. For example, Oliva and Goncalves [21] report that only 2% of their participants had heard of the beer game. But with increasing attention paid to the BWE and beer game in operation management courses in the past 20 years [15], more and more people involved in procurement and inventory management are familiar with the BWE. Second, the level of trust among the participants plays an essential role in the experiment’s outcome. If the participants do not know each other or are assigned to different roles and teams randomly, as is common in published experiments, it is unlikely that players will trust each other to make competent decisions. In fact, Croson et al. [13] report that this lack of trust can cause the BWE even when the demand is constant and known to all players. However, zero trust rarely occurs among suppliers and buyers in the real world, especially after months or years of repeated cooperation.

Despite these limitations, beer game experiments still provide valuable information. For example, they can suggest the form of players’ underlying order decision rules, which can then be quantified using regression analysis based on the data collected in the experiments. However, players’ behaviors can be quite different from one another. Each player tends to emphasize different components in his or her decision rule, and with different magnitudes, making it difficult to assess the prevalence of the BWE from beer game experiments alone. Therefore, we couple our beer game experiment with a simulation study to examine the impact of a wide range of ordering behaviors on the supply chain as a whole. Whereas our beer game experiments provide a description of the behaviors of a particular set of players, our simulation studies postulate a wider range of behaviors and determine the resulting effect on the pattern of order volatility in the supply chain.

Using our beer game and simulation experiments, we explore the behavioral causes of the BWE and RBWE. Previous beer game studies (e.g., [12, 25]) have suggested that the BWE is caused by demand uncertainty and an under-weighting of the supply line. Our results confirm these findings. We also find the inverse result: that supply uncertainty and over-weighting of the supply line can cause the RBWE.

To illustrate the RBWE, we extend the common metaphor of the supply chain as a string or whip, with the left-hand side representing upstream supply and the right-hand side
Figure 1: String vibrations (a) with no amplification, (b) with a vibration shock, and (c) with a fixed point.

representing downstream demand. Demand variability is represented as a vibration applied to the right end of the string. It is well known that a base-stock policy is optimal at each stage of a serial supply chain, and thus the BWE does not occur, if demands and purchase prices are stationary, upstream supply is infinite with a fixed lead time, and there is no fixed order cost [18]. In this case, demand vibrations are transmitted without modification up the string, as in Figure 1(a). It has been argued [25] that demand spikes act as shocks applied to the right end of the string, and that these shocks amplify as they move up the string, causing the BWE (Figure 1(b)).

If, instead, the left end of the string acts as a “fixed point” (that is, it is immovable), then vibrations will tend to dampen (rather than amplify) as they move up the string (Figure 1(c)). Such a fixed point may represent an upstream supply shortage: The upstream stage constantly produces at full capacity, so it has no variability in its order quantities. The vibration (demand) amplitude can only increase as one moves downstream—the RBWE.

Alternately, we can think of the RBWE a magnification of a wave that initiates upstream and propagates downstream (rather than as the dampening of a wave that propagates upstream). In this type of RBWE, a shock (in the form of a random capacity change) is applied to the upstream portion of the bullwhip, and the wave amplitude increases as it propagates downstream. We discuss both types of RBWE in this paper.

The remainder of the paper is organized as follows. In Section 2, we review the relevant literature. In Section 3, we explain the basic settings for our experiments and propose several order functions to model players’ ordering behavior. Sections 4 and 5 discuss the results of our beer game and simulation experiments, respectively. Finally, we summarize our conclusions in Section 6.
2. Literature Review

Traditional research on supply chain management (SCM) focuses on centralized decision-making problems such as facility location, logistics network design, multi-echelon inventory management, and so on. Under increasingly intense global competition, companies have recently tended to place more emphasis on their own core competencies. This trend has created a number of opportunities for supply chain coordination. Accordingly, an increasing amount of research effort has been devoted to decentralized supply chain management, including topics such as supply chain contracting and information sharing, and the BWE.

The theoretical study of the BWE was initiated by Lee, et al. [18], who introduced four operational causes for the BWE. Subsequently, a number of papers have explored these causes in greater detail. These papers are too numerous to review here but are reviewed in a recent article by Lee, et al. [17].

Most research in SCM tries to answer the question of what managers should do. In contrast, a recent focus on behavioral studies in operation management provides insight into what managers would do under various settings. Such research not only validates established theories but also encourages the development of new ones. Most behavioral studies of the BWE are carried out through the beer game. Sterman [25] introduces the game and observes order amplification due to the underweighting of the supply line—that is, players tend to ignore some or all of their pipeline inventory and instead base their ordering decisions primarily on their on-hand inventory.

A number of additional beer game experiments have attempted to confirm demand signal processing as a cause of the BWE [18], as well as other behavioral factors. Kaminsky and Simchi-Levi [15] find that a reduction in order information delay and shipment leadtime results in lower total supply chain costs but not in order variability amplification. Chen and Samroengraja [7] make the mean and standard deviation of the normally distributed demand known to every player and find that, though the four operational causes are removed, the BWE still occurs. Croson and Donohue [11] divide the participants into two groups. The participants in the control group only know the underlying demand distribution, while those in the POS treatment group know both the distribution and the realized customer demand. They observe a decrease in magnitude of the BWE in the POS treatment group compared with the control group. The primary reason for the decrease is that the participants in the POS treatment group almost equally utilize the realized customer demand and the order
information from their immediate downstream stage in their order decisions. Steckel et al. [24] show that POS information can actually increase the cost when it distracts the participants under certain types of customer demand.

Croson and Donohue [12] let the participants know the status of the inventory across the supply chain. The magnitude of the BWE decreases compared with the situation in which participants are not provided with such information, because upstream stages use downstream inventory information to anticipate and adjust their orders. Oliva and Goncalves [21] suggest that participants respond differently to their own on-hand inventory and backorders due to the difference between the holding and backorder costs and find that participants tend to ignore their own backorders rather than over-reacting to them and placing panicked orders. Wu and Katok [27] show that effective communication along with learning can significantly diminish the magnitude of BWE. Croson et al. [13] show that the BWE still persists even if the demand is constant and known to every player. They attribute this to “coordination risk”; that is, players place larger than necessary orders to protect themselves against the risk that other players will not behave optimally.

Despite its potential to systematically investigate the outcomes generated by various ordering behaviors, simulation has rarely been used in the BWE literature. Chatfield, et al. [6] use simulation to study the effect of players’ behaviors in a computer-only beer game. Their order functions are very similar to the base stock policy from Chen, et al. [8]. Using simulation, they find that an increase in the variance of the stochastic lead time results in greater BWE, while information sharing dampens BWE. Furthermore, they provide three forecast models under stochastic lead times. The model forecasting demand and leadtime separately leads to higher forecast variability and therefore higher order variability than the other two, one ignoring leadtime uncertainty and the other forecasting leadtime demand. The BWE reaches its maximum magnitude when demand and leadtime are estimated separately.

Another area of behavioral study in supply chain is the newsboy problem. Behavioral studies of the newsboy problem can provide valuable insight into managers’ actions in a more complex coordination mechanism, since the newsboy problem is one of the most important building blocks of supply chain management. Schweitzer and Cachon [22] study players’ responses to high-profit and low-profit products in the newsboy setting and find that orders are close to the demand mean regardless of the product type. This behavior can be explained by the anchoring and adjustment method [26], in which participants treat the mean demand as a starting point and make adjustments toward optimal solution. Bolton and
Katok [4] demonstrate that participants in a newsboy game tend to make decisions based on only partial data of the realized demand, and that the flat newsboy profit curve prevents participants from approaching the optimal order quantity quickly. Keser and Paleologo [16] extend the newsboy problem to the wholesale price contract and find that participants are willing split the profit roughly equally between them, despite the ability of one player to take the majority of the profit.

Readers interested in behavioral studies in other aspects of operation management are referred to the survey by Bendoly et al. [3]. They categorize such studies into three groups based on the assumptions used—intentions, actions, and reactions. Using this categorization, they review all of the behavioral papers that were published in six journals from 1985 to 2005.

3. Basic Settings and Order Decision Rules

In our beer game experiment and simulation, we study a 4-stage serial supply chain under periodic review. Stages 1–4 correspond to the retailer, wholesaler, distributor, and manufacturer, respectively. The retailer receives demand from an external customer. The demand follows a normal distribution with mean 50 and standard deviation (SD) 10. The manufacturer receives product from a single external supplier that is perfectly reliable and always had adequate supply. On the other hand, the manufacturer has an order capacity that limits the quantity it may order in a given period; see Section 3.1.

In each period, each stage $i$ experiences the following sequence of events:

1. The shipment from stage $i + 1$ shipped two periods ago arrives at stage $i$ (that is, the leadtime is 2). If $i = 4$, stage $i + 1$ refers to the external supplier.

2. The order placed by stage $i - 1$ in the current period arrives at stage $i$. If $i = 1$, stage $i - 1$ refers to the external customer.

3. Stage $i$ determines its order quantity and places its order to stage $i + 1$.

4. The order from stage $i - 1$ is satisfied using the current on hand inventory, and excess demands are backordered. Holding and/or stockout costs are incurred.

Unlike Sterman [25] and subsequent papers, we assume no order information delay, i.e., stage $i$ receives order information in the same period that the order is placed by stage $i - 1$. 
3.1 State Variables

We introduce the following random variables that describe the state of the system in any period:

- $IL_i^t$: Inventory level (on-hand inventory − backorders) at stage $i$ after event 2 (i.e., after observing its demand but before placing its order) in period $t$.
- $IP_i^t$: Inventory position (on-hand inventory + on-order inventory − backorders) at stage $i$ after event 2 in period $t$.
- $IT_i^t$: Inventory in transit to stage $i$ after event 2 in period $t$.
- $O_i^t$: Order quantity placed by stage $i$ in event 3 in period $t$. If $i = 0$, $O_i^t$ represents demand from the external customer.
- $\hat{O}_i^t$: Forecast of order quantity that will be placed by stage $i$ in period $t$. This forecast is calculated by stage $i + 1$ after event 2 in period $t - 1$. (See Section 3.2.4.)
- $\xi_t$: Order capacity at stage 4, i.e., the maximum quantity that stage 4 can order, in period $t$.
- $\sigma_t$: Standard deviation (SD) of orders placed by the stage $i$ across the time horizon.

Note that the order quantity $O_i^t$ is listed as a random variable, rather than as a decision variable, because it is determined by an order quantity function (see Section 3.2) whose terms are random state variables.

The order capacity at stage 4, $\xi_t$, is also a random variable. In our beer game experiment, the baseline capacity level was set to 70 (the demand mean plus two standard deviations). The beer game simulates disruptions by gradually reducing the capacity until it reaches 40; the system then recovers to full capacity. The timing of these disruptions is random to the players but is deterministic across teams so that all teams are playing with the same underlying data. See Section 4.1 for more details. Some of our simulation experiments also use randomly changing capacity, while others use a fixed capacity throughout the time horizon.
3.2 Order Quantity Functions

We introduce three order quantity functions that express the order quantity as a function of several state variables. The base order function comes directly from Sterman [25]. The extended order function is used to address the fact that players may separate their in-transit inventory from their on-order inventory; that is, they may consider backorders at the supplier separately from in-transit inventory. The capacity shock order function allows players to behave differently based on whether the capacity is in its normal or disrupted state.

In our simulation studies, these functions are used to determine the order quantity placed by each stage in each period. In our beer game experiment, we calibrate players’ observed behaviors to these functions to determine which state variables players consider when choosing an order quantity, and to what extent. This analysis is motivated by that of Sterman [25].

3.2.1 Base Order Function

Our base order function is identical to the function proposed by Sterman [25]:

$$O_t^i = \max\{0, \hat{O}_{t+1}^i + \alpha_s^i(IL_t^i - a_s^i) + \beta_s^i(IP_t^i - IL_t^i - b_s^i)\},$$

where $a_s$ and $b_s$ represent target values for the inventory level ($IL$) and supply-line inventory ($IP - IL$). The constants $\alpha_s$ and $\beta_s$ are adjustment parameters when the actual inventory level and supply line, respectively, deviate from the desired targets. (The subscript $s$ stands for “Sterman.”) Sterman based this order function on the anchoring and adjustment method proposed by Tversky and Kahneman [26]. It accounts for changes in the demand, inventory level, and supply line dynamically, even when the demand and supply processes are unknown. $\hat{O}_{t+1}^i$ is treated as the anchor, serving as a starting point for the order quantity, while the remaining part is the adjustment to correct the initial decision based on the inventory level and supply line.

The order quantity placed by stage 4 in period $t$ is bounded by its capacity $\xi_t$ in that period. Therefore the actual order placed by stage 4 in period $t$ is $\min\{O_t^4, \xi_t\}$.

3.2.2 Extended Order Function

In the traditional beer game setting, each stage is assigned a per-unit holding and shortage cost. Because participants know that their success is measured by the total cost of the team rather than on individual players’ performance, participants may take into consideration
the backorders at their suppliers and treat these differently from in-transit inventory when making order decisions. That is, a player may order less than his or her desired quantity if the supplier stage is experiencing backorders, since those backorders contribute to the team’s cost.

Oliva and Goncalves [21] find that participants often ignore their own backorders and only respond to their on-hand inventory (that is, they ignore the negative part of $IL$ and only use the positive part to decide an order quantity). One explanation for this behavior is that the participants are willing to incur backorders themselves to reduce the backlog at their suppliers. Indeed, during the post-game questionnaire that we gave to our beer game participants, several participants reported they would “decrease [their] order” and not “push the supplier too much” when their suppliers fail to deliver.

However, the base order function (1) does not distinguish between on-order items that are in transit and those that are backordered. Therefore, we propose an extended model in which the supply line is split into in-transit inventory and backorders at the supplier:

$$O^i_t = \max\{0, \hat{O}^i_{t+1} + \alpha^i_e(\hat{IL}^i_t - a^i_e) + \beta^i_e(IT^i_t - b^i_e) + \gamma^i_e(IP^i_t - IL^i_t - IT^i_t - c^i_e)\},$$  

(2)

where $a^i_e, b^i_e$ and $c^i_e$ are target values for the inventory level ($IL$), in-transit inventory ($IT$), and backorders at the upstream stage ($IP - IL - IT$). The constants $\alpha^i_e, \beta^i_e,$ and $\gamma^i_e$ are adjustment parameters when actual inventory level, in-transit inventory, and upstream backorders, respectively, deviate from the desired targets. (The subscript $e$ stands for “extended.”) Note that $IP^i_t - IL^i_t - IT^i_t$ is always nonnegative. If $\gamma^i_e < 0$, then stage $i$ is willing to order less when stage $i+1$ has backorders to alleviate the backorder cost at stage $i+1$. If $\gamma^i_e = 0$, then the stage ignores upstream backorders. If $\gamma^i_e > 0$, then stage $i$ orders more when stage $i+1$ has backorders. This may happen if multiple stages compete for scarce resources in a shortage game (which is not included in the traditional beer game), or if the player finds that the upstream stage consistently cannot fulfill its order and thus inflates the order to encourage the supplier to keep adequate stock on hand in the future.

For stage 4, the extended order function degenerates to the base order function since the outside supplier always satisfies the order placed by the stage 4 in beer game.

### 3.2.3 Capacity Shock Order Function

In reality, a huge capacity shock is usually known publicly. This information about capacity shocks may have an impact on buyers’ behavior, e.g., by increasing purchase quantities. In
order to examine whether players behave differently with or without capacity shocks, we propose the capacity shock order function:

\[
O^i_t = \max\{0, \hat{O}^{i-1}_{t+1} + \alpha_d^i(IL^i_t - a^i_d) + \beta_d^i(\mathcal{I}P^i_t - IL^i_t - b^i_d) + \gamma_d^i S_t\},
\]

where \(S_t\) is a public signal to indicate whether there is a capacity shock in the system. \(S_t = 1\) if there is capacity shock and 0 otherwise. (The subscript \(d\) stands for “disruption.”)

If \(\gamma_d^i < 0\), then the stage is willing to order less during a capacity shock in order to reduce potential backorders at its supplier. If \(\gamma_d^i = 0\), the stage ignores capacity shocks, while if \(\gamma_d^i > 0\), then the stage orders more during a capacity shock to protect against further decreases in capacity.

As in the base order function, the actual order quantity placed by stage 4 is given by \(\min\{O^4_t, \xi_t\}\).

### 3.2.4 Demand Forecast

In each of the three order functions, \(\hat{O}^{i-1}_{t+1}\) is a forecast of demand that will be observed by stage \(i\) in period \(t + 1\). It is calculated using exponential smoothing:

\[
\hat{O}^{i-1}_{t+1} = \eta O^{i-1}_t + (1 - \eta)\hat{O}^{i-1}_t,
\]

where \(\eta\) is the smoothing factor.

### 3.3 Interpretation of Order Standard Deviation

We examine the presence of the BWE or RBWE at each stage individually, as well as in the supply chain as a whole. When \(\sigma_i > \sigma_{i-1}\), stage \(i\) amplifies its order variability; i.e., the bullwhip effect (BWE) occurs at stage \(i\). If \(\sigma_i < \sigma_{i-1}\), then the reverse bullwhip effect (RBWE) occurs at stage \(i\) instead.

If \(\sigma_{i+1} > \sigma_i\) for all \(i \leq 3\), then the system exhibits pure BWE, and when \(\sigma_{i+1} < \sigma_i\) for all \(i \leq 3\), the system exhibits pure RBWE. There are also several other possible shapes for the order standard deviation pattern across the supply chain. For example, when \(\sigma_{i+1} > \sigma_i\) for \(i = 1, 2\) and \(\sigma_{i+1} < \sigma_i\) for \(i = 3\), the order pattern resembles an umbrella; this pattern is natural when the downstream part of the supply chain is affected primarily by demand uncertainty while the upstream is bounded by the capacity process.
4. Beer Game Experiment

In our beer game experiment, we create capacity shocks during the game to observe how players behave during supply disruptions. All players know when a capacity shock is occurring, but only the player in the role of the manufacturer knows its severity. We expect some players to overreact to this supply uncertainty by inflating their orders, leading to greater order variability. This overreaction may cause the RBWE rather than the BWE. Our goal is to test how players respond under the pressure of uncertainty from both the demand side and the supply side.

Our beer game setup is motivated in part by consumer buying patterns for gasoline following hurricane Katrina. Katrina disabled approximately 10% of U.S. oil refining capability, as well as a substantial portion of its drilling facilities [19], creating a sudden and severe shock to the upstream stages of the gasoline supply chain. Customers were aware of this shock (but not of the magnitude of its downstream effect), and many of them filled their cars at the beginning of the shock in order to avoid future shortage and price fluctuations. Although the impact of price fluctuations is not tested in the beer game, we can still test players’ response to possible shortages created by the supply uncertainty.

In our experimental results, we partially observe the RBWE created by supply uncertainty during disruptions. But, contrary to our hypothesis that the RBWE occurs because players over-order during disruptions, we find that players order less during disruptions. We believe this is because the beer game is centralized, in that players are evaluated by the performance of the whole team. This setting is different from the Katrina example, in which customers act in their own best interests.

4.1 Experimental Design

Our beer game experiment was conducted using an Excel-based implementation written by the authors. Our computerized implementation gives players more information about the status of the system than in the traditional board version of the game. Figure 2 shows the game’s user interface. Players can easily acquire information about their own on-hand inventory, backorders, on-order inventory, and in-transit inventory, as well as backorders at their supplier. The game also provides players with information about their expected inventory level in the next two periods (using their current in-transit inventory and assuming the demand will equal its mean). This information reminds participants about their in-transit
inventory and helps them understand their supply line better.

No communication is allowed during the game. Our implementation automates the information-transfer process: when a player places an order, it is transmitted electronically to his or her upstream neighbor, and when orders are shipped, the delivery quantity is transmitted downstream electronically. This reduces transaction errors and speeds the playing of the game.

Following Sterman [25], we set the holding and backorder cost to $.5 and $1, respectively, at every stage of the supply chain. Demands appear random to the players but were generated prior to the game from a $N(50, 100)$ distribution so that each team would play with the same data.

The manufacturer has a capacity limitation on his or her order size. The capacity fluctuates throughout the game, starting at 70 (mean demand plus two standard deviations), then gradually decreasing to 40, then returning to 70. Players were informed of this rough pattern, but the actual timing of the changes was unknown to them. Table 1 describes the capacity process in detail; both teams faced the same pattern of changes. Players were notified (via an indicator in the beer game screen) whether the capacity is full or short at each period, but they did not know how severe the capacity shortage was. The exception is the manufacturer, who can determine the capacity in any period since the program prompts
Table 1: Capacity process of the beer game

| Period | 1-5 | 6-10 | 11-15 | 16-20 | 21-28 | 29-36 | 37-40 | 41-46 | 47- |
|--------|-----|------|-------|-------|-------|-------|-------|-------|-----|
| Capacity | 70  | 60   | 50    | 40    | 70    | 60    | 50    | 40    | 70  |

him or her for a new order quantity if the quantity entered exceeds the capacity.

Our experiment consisted of eight participants (two teams of four) from the ISE department at Lehigh University, including seven Ph.D. students and one professor. The players were quite familiar with each other, eliminating some of the trust issues present in other games. Only two players had never heard of the BWE before, and most had studied it in SCM classes. Players received a monetary incentive to encourage them both to participate and to play to the best of their abilities. The incentive ranged from $5 to $20 and was based on the teams’ relative performance (total average cost per period), following the allocation rule from Croson and Donohue [11].

4.2 Beer Game Results

4.2.1 Overall Results

The results from the beer game experiment are summarized graphically in Figure 3. Descriptive statistics are given in Table 2, which lists statistics both overall and for groups of periods with similar capacity patterns: 1–5 (full capacity), 6–20 (capacity shock), 21–28 (full capacity), 29–46 (capacity shock). Note that team 1 played for 49 periods, while team 2 only played to period 39 in the time allotted.

We see two different patterns of order standard deviation in Table 2. In team 1, the order SDs of the first three players are nearly equal, and there is a sudden drop in SD at the manufacturer due to the capacity constraint. In the second team, players 1–3 exhibit RBWE, and player 4 (manufacturer) displays an increase in SD, which may be due to the capacity shock (since variability in capacity results in variability in orders). Further support for the existence of the RBWE will come from our rejection of Hypothesis 2 below and from Section 5.6, which together imply that, in the presence of capacity shocks, the RBWE can occur even when players underweight the supply line (which otherwise tends to cause BWE).

The retailer, wholesaler, and distributor (stages 1–3) all demonstrate oscillating patterns in their order behaviors. For the winning team (team 2), the amplitude of oscillation goes
Table 2: Results from Beer Game Experiment

| Role | Overall | 1-5 | 6-20 | 21-28 | 29-46 | Overall | 1-5 | 6-20 | 21-28 | 29-46 |
|------|---------|-----|------|-------|-------|---------|-----|------|-------|-------|
| R1   | 47.74   | 40.00| 48.13| 59.88 | 51.33 | 21.51   | 13.23| 8.75 | 4.88  | 27.89 |
| W1   | 47.20   | 42.00| 48.67| 60.50 | 51.33 | 22.74   | 10.95| 10.77| 5.86  | 27.89 |
| D1   | 52.9    | 25   | 59.33| 70    | 58.89 | 21.69   | 12.25| 11.16| 0     | 13.35 |
| M1   | 52.56   | 30.00| 47.87| 70.00 | 51.11 | 14.66   | 21.79| 7.05 | 0.00  | 9.00  |
| R2   | 51.13   | 56.00| 49.67| 47.75 | 53.36 | 7.83    | 8.94 | 8.34 | 5.95  | 7.02  |
| W2   | 51.26   | 58.00| 49.00| 48.00 | 53.64 | 6.32    | 5.70 | 7.37 | 4.41  | 1.96  |
| D2   | 51.51   | 58.60| 49.13| 49.50 | 53.00 | 5.00    | 5.90 | 4.49 | 1.93  | 3.32  |
| M2   | 51.82   | 68.20| 43.00| 61.88 | 49.09 | 12.35   | 2.05 | 8.19 | 10.33 | 8.31  |

*aColumn headers indicate ranges of periods.*
down by time, suggesting that some players improved their behavior over the course of the experiment.

4.2.2 Validity of Extended Order Function

To test whether players take backorders at their supplier into consideration and treat it differently from the in-transit inventory, as suggested in Section 3.2.2, we formulate the following null hypothesis:

**Hypothesis 1** Players treat backorders at their supplier the same as in-transit inventory. That is, \( \beta_e = \gamma_e \), and the extended order function degenerates to the base order function.

The statistical procedure used in previous beer game studies (e.g., [11, 12, 21]) to calibrate the order quantity function to observed data treat \(-\alpha_s a_s - \beta_s b_s\) as a single constant to be estimated. This leaves \(\alpha_s IL + \beta_s (IP - IL)\), and since the magnitude of the supply line \((IP - IL)\) is usually larger than the inventory level \((IL)\), this strategy can underestimate the weight placed on the supply line. A more accurate method would perform the regression on the differences \(IL - a\) and \(IP - IL - b\), since the magnitudes of the differences can be expected to be roughly equal. Unfortunately, estimating \(\alpha_s, \beta_s, a_s,\) and \(b_s\) as four separate constants would require non-linear regression.

In our analysis, we fix \(b_s\) (and \(b_e\) for the extended order function) to be 100, since the participants were told that the mean demand is 50 units per period and the leadtime is 2. In addition, we set \(\eta = 0\) since the demand process is known and no forecasting is necessary. For the purposes of our regression, we can assume that participants use the mean demand as the anchor and account for the actual observed demand through the inventory level term in the order function.

We vary \(a_s\) (\(a_e\)) from 0 to 20 by increments of 1. For each value of \(a_s\) (\(a_e\)), we apply linear regression (without fitting the intercept) to both the base and extended order functions and choose the value of \(a_s\) (\(a_e\)) with the least sum of squares due to error (SSE). We use F statistics and p-values to test Hypothesis 1. We do so with both individual and aggregated data. Since the base and extended order functions are equivalent for the manufacturer, we exclude the manufacturer data from the aggregate data. The results are shown in Table 3. (The last two rows refer to the capacity shock order function and will be used to test Hypothesis 2 below.) From the aggregate perspective, Hypothesis 1 should be rejected \((p_e < 0.0001)\). From an individual perspective, there are two participants (R1 and D1) for
Table 3: Regression Results

|       | R1 | W1 | D1 | R2 | W2 | D2 | Aggregation |
|-------|----|----|----|----|----|----|-------------|
| \(\hat{\alpha}_s\) | 0  | -0.0604 | -0.4022 | -0.1902 | -0.3458 | -0.1970 | 0.1939 |
| \(\hat{\beta}_s\) | 0.0643 | -0.2806 | -0.0286 | -0.0937 | 0.0347 | -0.0468 | -0.0090 |
| \(\hat{\alpha}_e\) | 0  | 0  | -0.757 | -0.3778 | -0.1602 | -0.3323 | -0.1983 | -0.2364 | -0.1105 |
| \(\hat{\beta}_e\) | 0.3359 | -0.2314 | 0.1874 | -0.1121 | 0.0759 | -0.0621 | 0.2152 |
| \(\hat{\gamma}_e\) | -0.0128 | -0.2640 | -0.0311 | -0.0459 | -0.0100 | -0.0302 | -0.0221 |
| \(\hat{\alpha}_d\) | 0  | 0  | -0.1157 | -0.3744 | -0.2094 | -0.3749 | -0.2010 | -0.2299 | -0.1515 |
| \(\hat{\beta}_d\) | 0.1610 | -0.1864 | -0.0197 | -0.1140 | 0.0384 | -0.0295 | 0.00752 |
| \(\gamma_d\) | -22.1178 | -13.9210 | -5.0819 | -2.2193 | -0.7558 | 0.7968 | -4.6265 |
| \(F_e\) | 20.6166 | 0.5385 | 3.7590 | 0.2801 | 0.8417 | 0.5231 | 48.7896 |

| \(p_e\)-value | <0.0001 | 0.4668 | 0.0587 | 0.5999 | 0.3650 | 0.4742 | <0.0001 |
| \(F_d\) | 33.0034 | 9.6067 | 2.0843 | 0.8248 | 0.2668 | 0.1929 | 14.7564 |

| \(p_d\)-value | <0.0001 | 0.0033 | 0.1556 | 0.3698 | 0.6087 | 0.6632 | 0.0002 |

whom Hypothesis 1 should be rejected. The reason for the high \(p_e\)-values in team 2 may be the relatively smaller team-wide backorders compared with team 1 (and therefore fewer data points to test how players react to backorders). Notice that all the signs of \(\hat{\gamma}_e\) are negative, suggesting that participants do indeed treat upstream backorders separately from on-order quantities. In addition, \(\hat{\alpha}_s < \hat{\beta}_s\) for every player, confirming that players under-weight the supply line.

4.2.3 Validity of Capacity Shock Order Function

To test whether people behave differently during a capacity shock, we test the following null hypothesis:

**Hypothesis 2** Players ignore the capacity shock signal. That is, \(\gamma_d = 0\), and the capacity shock order function degenerates to the base order function.

We use the same statistical procedure as in Section 4.2.2, except that we use the capacity shock order function in place of the extended order function. Since the manufacturer’s order quantity is bounded by the capacity limit, we exclude manufacturer data from the aggregate data. \(\hat{\gamma}_d\) From the last two rows of Table 3, we reject Hypothesis 2 for the aggregate data and for two players from team 1. The coefficient \(\hat{\gamma}_d\) is negative for all players except one (D2).
This suggests that players are willing to decrease their intended demand during capacity shocks to prevent backorders at their suppliers. Another interesting finding is that $|\hat{\gamma}_d|$ and its corresponding $p_d$-value is largest at the retailers. This suggests that the most downstream participants react the most strongly to capacity shocks. In addition, team 1 responded more to capacity shocks than team 2 did, and this difference hurt team 1’s overall performance and resulted in higher average costs.

From our rejection of Hypotheses 1 and 2 in the aggregate and for certain players, and from the signs of $\hat{\gamma}_e$ and $\hat{\gamma}_d$, we conclude that participants try to reduce backorders at their suppliers, especially during disruptions. This reaction to capacity shocks may be the cause of the RBWE even though players still underweight the supply line, as $\hat{\alpha}_d$ is still more negative than $\hat{\beta}_d$.

5. Simulation

There are two objectives in our analysis of the data from the beer game experiment. One is to examine the behavior of individual players and the other is to make inferences about ordering decisions across all the participants. Inferences based on all participants provide a picture of how players behave in general. However, the total cost of each supply chain is highly dependent on the behavior of the individuals in the chain, rather than on the general behavior. It is the individual behaviors and the arrangement of the players in a team that matter. Additionally, from previous studies [12, 25], we know that individuals behave quite differently from each other. The behaviors of the participants in a given team may interact strongly. The same player, if assigned to different groups, may even behave differently due to the effect of other players.

Another drawback of the beer game experiment is the time limit. To achieve some sort of stable behavior, the participants need to learn the order patterns of their customer and supplier, and this may take a long time. This suggests that the parameters in the estimated order function vary over time at the beginning of the horizon. However, the time limitation prevents the beer game from being played long enough to achieve stability.

In addition, there are differences between the beer game setting and a real business setting. In the beer game, the total cost of the supply chain is the performance measure, while real businesses care about their own profit, not (directly) that of their partners. This may cause different values of the parameters in the order function, e.g., the customers in the
Katrina example may have positive $\gamma_d$ instead of negative ones. Simulation can be easily adapted to the various settings by changing the values of the parameters.

For these reasons, we performed an extensive simulation study to address a wide variety of possible human behaviors and to examine the impact of players’ behavior once it reaches an approximately steady state. Our simulation study complements our beer game experiment.

We used the freeware software called BaseStockSim developed by Snyder [23], which simulates multi-echelon supply chains with stochastic supply and demand. The user can choose among several types of inventory policies for each stage, including base stock, $(r, Q)$, $(s, S)$, and various anchoring and adjustment order functions.

Except where indicated, the capacity at stage 4, $\xi_t$, is infinite. In Sections 5.1–subsec:individual, we use the base order function. We also simulated the system using the extended order function, but the results were quite similar, so we omit them here. In Section 5.6, we allow the capacity to change randomly and use the capacity shock order function. Except in Sections 5.5 and 5.6, all stages follow the same order function as each other, with the same parameters.

Table 4 shows our baseline values for the parameters in the base order function. Our simulation experiments use these settings as a baseline and then modify them individually to test the effect of changes in various parameters. The baseline values are based on the experimental results of Sterman [25], with slight modifications. We keep the structure of underweighting the supply line ($\beta_s < 0$). In the simulation study, we limit $-1 \leq \alpha_s^i, \beta_s^i \leq 0$ and $0 \leq \eta_s^i \leq 1$. The bounds prevent stages from reacting to deviations in the inventory level and supply line more than the demand and eliminate potential instability in the system.

| Parameter | Base Value |
|-----------|------------|
| $\alpha_s^i$ | $-0.3$ |
| $\beta_s^i$ | $-0.2$ |
| $\eta_s^i$ | $0.3$ |
| $a_s^i$ | $14$ |
| $b_s^i$ | $100$ |

Our goal is to simulate the system with an extensive range of parameter values to determine the impact of various behaviors on ordering patterns across the whole system. For each setting of the parameters, we simulated the system for 10 trials, each consisting of 1000
Figure 4: Effect of $\alpha_s$ ((a) and (b)) and $\beta_s$ ((c) and (d)) under the base order function.

5.1 Effect of Weights on Inventory Level and Supply Line

Figure 4 shows the order standard deviation pattern across the supply chain under the base order function as $\alpha_s$ and $\beta_s$ change from the base case separately. The closer $\alpha_s$ is to $-1$, the greater the SD of orders at all stages is in general. The order SD is a bowl-shaped function of $\beta_s$: it initially decreases with $\beta_s$, then increases. Figures 4(b) and (d) provide greater magnification of the curves in Figures 4(a) and (c).

The majority of instances in these Figures 4(b) and (d) exhibit pure RBWE (largest SD at stage 1, smallest at stage 4), and the curves tend to cross when $\alpha_s = \beta_s$. This leads us to conjecture that pure RBWE tends to occur when the weight placed on the inventory level is less negative than that placed on the supply line ($|\alpha_s| < |\beta_s|$), and that pure BWE tends to occur when the opposite is true ($|\alpha_s| > |\beta_s|$). In other words, the system behaves differently depending on whether stages account more for today’s costs (inventory level) or future costs (supply line) when choosing order quantities.

To examine whether this conjecture holds more generally, we divide the parameter space into a grid and simulate the system at all of the grid points. The grid settings without capacity limits are shown in the left-hand portion of Table 5. (The right-hand portion will
Table 5: Grid Settings of Parameters for Base Order Function

| Parameter | Without Capacity Limitation | With Capacity Limitation |
|-----------|-----------------------------|--------------------------|
| $\alpha_s$ | 0, -0.2, -0.4, -0.6, -0.8 | 0, -0.2, -0.4, -0.6, -0.8 |
| $\beta_s$ | 0, -0.2, -0.4, -0.6, -0.8 | 0, -0.2, -0.4, -0.6, -0.8 |
| $\eta_s$ | 0.1, 0.3, 0.5 | 0.1, 0.3, 0.5 |
| $a_s$ | 0, 14, 28 | 0, 14, 28 |
| $b_s$ | 100, 114 | 100, 114 |
| capacity | $\infty$ | 55 |

Table 6: SD Patterns from Grid Experiment Without Capacity Constraint

| Type | Total | Percent | $|\alpha_s| > |\beta_s|$ | $|\alpha_s| = |\beta_s|$ | $|\alpha_s| < |\beta_s|$ |
|------|-------|---------|-----------------|-----------------|-----------------|
| 1234 (BWE) | 231 | 51.33% | 180 | 30 | 21 |
| 4321 (RBWE) | 171 | 38.00% | 0 | 18 | 153 |
| 3421 | 13 | 2.89% | 0 | 12 | 1 |
| 2314 | 12 | 2.67% | 0 | 12 | 0 |

be used in Section 5.4.) These settings result in a total of 450 instances simulated.

Table 6 summarizes the results of these experiments. The entry in the first column indicates the ordering of the SD for the four stages, from smallest to largest. For example, 1234 indicates that stage 1 has the smallest order SD, stage 2 has the second smallest, and so on. The pattern 1234 represents pure BWE and 4321 represents pure RBWE. We observed 8 types of order SD patterns in our results. Table 6 only lists the patterns that occurred in more than 2% of the 450 instances tested.

Out of the 450 instances, $|\alpha_s| < |\beta_s|$ in 180, $|\alpha_s| > |\beta_s|$ in 180, and $|\alpha_s| = |\beta_s|$ in 90. The three right-most columns of Table 6 indicate the number of instances (out of those counted in column 2) for which each case holds. From the table we can see that the BWE occurs whenever $|\alpha_s| > |\beta_s|$, and pure RBWE occurs nearly 85% of the time when $|\alpha_s| < |\beta_s|$. This supports our conjecture that BWE tends to occur when $|\alpha_s| > |\beta_s|$ and RBWE tends to occur when the opposite is true.

If the inventory level is weighted more than the supply line, we get pure BWE for sure. On the other hand, if the inventory level is weighted less heavily than the supply line, pure RBWE is likely but not guaranteed. When the two elements are equally weighted, the pattern of order SDs across the supply chain is quite uncertain.

21
5.2 Effect of Target Levels

We vary the target values of the inventory level and supply line under the base order function individually from the base setting. The results are shown in Figure 5.

As $a_s$ (target inventory level) increases, so does the stage’s safety stock. From Figure 5(a), order SDs decrease across the supply chain as $a_s$ increases. In other words, inventory can reduce the magnitude of the BWE. As inventory levels increase, so do service levels; panicky ordering is reduced as a result, which reduces order SDs and further improves service levels upstream. This self-reinforcing cycle is the mechanism by which BWE is reduced when $a_s$ increases. On the other hand, the benefit from reducing BWE in this way can be outweighed by the cost of the additional inventory.

Similar results hold as $b_s$ (target supply line) increases, since increasing $b_s$ in (1) is equivalent to increasing $a_s$.

5.3 Effect of Exponential Smoothing Factor

In this section, we vary the exponential smoothing factor to determine its effect on the SD pattern. The results are shown in Figure 6. The magnitude of BWE at all the stages increases with the exponential smoothing factor, since larger values of $\eta$ result in less smooth forecasts and therefore more variable orders. These results confirm the findings of Chen et al. [8, 9]. Fortunately, managers tend to use small exponential smoothing factors in practice.

5.4 Effect of Finite Capacity

The effect of changes in capacity under the base order function is shown in Figure 7, which displays the order SD at each stage as a function of the capacity at stage 4, which is constant.
over time throughout the simulation. The capacity has a tremendous impact on the order SD pattern at stage 4. The capacity serves as an upper bound for the order quantity for stage 4, which is given by \( \min\{\xi_t, O_t^4\} \). No matter how volatile the orders placed by stage 3 are, stage 4 cannot respond freely due to the capacity limit. It is not surprising that the tighter the capacity is, the smoother the order pattern of stage 4 becomes. When stage 4 has tight enough capacity (\( \leq 54 \)), the supply chain as a whole will exhibit RBWE in the sense that \( \sigma_4 < \sigma_1 \), although stages 1–3 individually exhibit BWE.

In Figure 8, we plot the orders of all four stages from period 501 to period 600 with \( \xi = \infty \) (part (a)) and \( \xi = 55 \) (part (b)). There is a clear oscillating pattern in both cases, although the order function is linear. Cachon et al. [5] show that the RBWE is likely to occur when the demand is seasonal. One underlying reason for this phenomenon may be that the capacity cuts the peak of the demand oscillation. Since there is not enough capacity to supply the peak of the demand, the orders placed by the stage 4 can only reach the capacity limit, not the actual demand requirement.

When there is a capacity limit at stage 4, the amplitude of the order cycle decreases at all stages. The orders in Figure 8(a) range from 0 to 120, while those orders in Figure 8(b)
range from 20 to 100. One possible reason is that when the capacity at stage 4 is tight, the supply line inventory is more stable. The closer a stage is to the outside supplier, the more stable the supply line is. Therefore some of the uncertainty in the supply line is removed, and stages primarily need to deal with demand uncertainty. In Figure 7, it is evident that when the capacity is finite, the order SD decreases more significantly for stages that are closer to stage 4.

We also use a grid evaluation to test the impact of capacity under a broader range of instances. The grid settings are given in the right-hand portion of Table 5, and the results are given in Table 7. The results exhibited 8 distinct patterns of order SD. Again, only patterns occurring in at least 2% of the instances are included in the table. Fewer than 2% of the instances exhibit pure BWE. In more than 87% of the instances, stage 4 has the smallest order SD, suggesting RBWE in at least a portion of the supply chain. Patterns 4123 and 1423 occur when $|\alpha_s| > |\beta_s|$; in these patterns, the BWE occurs at stages 1–3. Meanwhile, the pure RBWE occurs in over 40% of the instances. This suggests that even if managers behave suboptimally in a supply chain, tight capacity prevents severe BWE from occurring.

Table 7: Grid Experiment under Base Order Function with Capacity Constraint

| Type       | Total | Percent | $|\alpha_s| > |\beta_s|$ | $|\alpha_s| = |\beta_s|$ | $|\alpha_s| < |\beta_s|$ |
|------------|-------|---------|-----------------|-----------------|-----------------|
| 4321 (RBWE)| 185   | 41.11%  | 0               | 36              | 149             |
| 4123       | 175   | 38.89%  | 127             | 30              | 18              |
| 1423       | 53    | 11.78%  | 53              | 0               | 0               |
| 4231       | 14    | 3.11%   | 0               | 14              | 0               |
| 4312       | 10    | 2.22%   | 0               | 0               | 10              |

Figure 8: Order Pattern under Base Order Function
5.5 Effect of Individual Behaviors

In the preceding simulation studies, all stages use the same parameters in the order function. This helps us to develop insight about the effects of various behaviors. But in reality, each stage may use different values for the parameters. Therefore, in this section we study the effect of an individual following an order behavior that differs from the rest of the supply chain. We only discuss the base order function. (Results using the extended order function were similar.)

In Figure 9, we change $\alpha_s$ (left-hand figures) and $\beta_s$ (right-hand figures) individually for each stage, keeping the parameters at the other stages set to the baseline values. For example, the top-left graph in Figure 9 plots the order SDs for each stage if stages 2–4 follow the baseline parameters while stage 1 uses a different value of $\alpha_s$. We refer to the stage following different parameters as “irrational” when it over-weights the inventory level or under-weights the supply line in the order function compared to the other stages.

For the most part, these figures exhibit pure BWE. This is not surprising since capacity is infinite. What is surprising is that, regardless of how $\alpha_s$ and $\beta_s$ change, the ordering behavior at stage $i + 1$ has only a minor impact on that of stage $i$. For example, as $\alpha_s$ increases at stage 2, the order SDs at stages 2–4 decrease, but that for stage 1 stays constant. Similarly, as $\alpha_s$ increases at stage 3, the order SD stays constant at stages 1 and 2, and so on. This is surprising since one would expect that poor ordering decisions at stage $i + 1$ will have an effect on downstream stages. No intuitive explanation is readily apparent to explain this phenomenon. Further study is required to determine whether it is a general trend or the result of our particular assumptions.

Examining the y-axis throughout Figure 9, we find that the order volatility caused by the irrational player is greatest when that player is at stage 1. From these two findings, we argue that the impact of demand uncertainty caused by irrational order behavior is more serious than the supply uncertainty brought about by irrational order behavior. The order variability of the supply chain depends both on individual behaviors and on how the players are located in the supply chain.

5.6 Effect of Supply Uncertainty

In this section, we incorporate supply shocks into our simulation model to confirm our findings from the beer game experiment that capacity shocks may cause the RBWE. We use
Figure 9: Effect of Individual Behavior on the Supply Chain
the capacity shock order function (Section 3.2.3) to model players’ reactions to disruptions. Given the long lines of end-customers waiting for gasoline after Katrina, and our beer game results indicating that the player furthest downstream reacts most strongly to disruptions, we set $\gamma_d \neq 0$ at stage 1 only; that is, only stage 1 changes its order quantity in response to disruptions.

We test two different values for $\alpha_d$ at every stage: $\alpha_d = -0.3$ (the baseline value) and $\alpha_d = -0.2$. The first value represents underweighting of the supply line while the second represents equal weight placed on the inventory level and the supply line. In both cases, we vary $\gamma_d$ to model stage 1’s reaction to disruptions.

We fix the demand at 50 per period in order to remove demand uncertainty from the model and focus only on supply uncertainty. The supply process follows a two-state discrete-time Markov process. The “up” state corresponds to full capacity $\xi_u$ and the “down” state corresponds to disrupted capacity $\xi_d$. The transition probability from the up state to the down state is $p_d$, and that from the down state to the up state is $p_u$. The stationary probabilities of being in the up and down state are $p_u/(p_u + p_d)$ and $p_d/(p_u + p_d)$, respectively.

To ensure the stability of the system, we require $(\xi_u p_u + \xi_d p_d)/(p_u + p_d) > 50$. That is, the service rate must be greater than the arrival rate. We test two values of $p_u$: 0.5 and 0.9.

Even though the capacity distribution is not iid (since the state in period $t$ depends on that in period $t-1$), we can still compute its long-run variance: The variance of the capacity is given by

$$(\xi_u - \xi_d) \sqrt{\frac{p_u p_d}{p_u + p_d}}.$$

The order placed by stage 4 is bounded above by $\xi_d$ or $\xi_u$ (depending on the state). We assume $\xi_d < 50 < \xi_u$.

In Figure 10, we set $\xi_u = 60$, $\xi_d = 40$ and $p_d = 0.1$. Due to stage 1’s reaction to capacity shocks, even underweighting of the supply line fails to generate BWE from stage 1 to stage 3 consistently. This is different from the results in Table 6 and 7, in which underweighting of the supply line always generates BWE at stages without a capacity limit. Notice that, in Figure 10(c), even if no stage reacts to disruptions ($\gamma_d = 0$), there is still order variability at all stages; i.e., the capacity shock transfers downstream. When the whole supply chain places equal weight on the inventory level and supply line, stages 1–3 exhibit BWE, while stage 4 exhibits RBWE.

When $p_u = 0.9$, the standard deviation of the available capacity is 6. When $p_u = 0.5$, the standard deviation of available capacity is 7. Since the available capacity is smaller than
the demand during a capacity shock, the large fluctuations in order quantity at stage 4 are mainly caused by fluctuations in the random capacity.

In Figure 11, we set $\gamma_d = -10$. In parts (a) and (b), we set $p_u = 0.9$ and vary $p_d$, while in parts (c) and (d), we set $p_d = 0.1$ and vary $p_u$. When disruptions are more frequent (parts (a) and (b)), stage 1 behaves in a more “panicky” manner, but this has a very small impact on stages 3 and 4. When disruptions take longer to recover (parts (c) and (d)), stage 3 exhibits “panicky” behavior, but the impact on stage 1 is quite small. This suggests that the stages close to the capacity shock are affected more by the recovery process, while the stage reacting to the disruption (stage 1) is affected more by the failure process. This may be because, the smaller the recovery rate is (assuming it is less than the failure rate), the larger the variance of the capacity is. The stages close to stage 4 are affected more by the supply uncertainty. At the same time, each stage carries some inventory, which may buffer the downstream against this uncertainty. Therefore, the upstream stages have more exposure to the supply uncertainty. On the other hand, the more frequently disruptions occur, the more frequently stage 1 responds. However, due to the self-correcting order quantity mechanism, the reaction to the disruption is mitigated by larger backorder quantities during the subsequent recovery process. So stage 1 is less affected by the recovery process. This is also similar to gas-buying patterns after hurricane Katrina. Customers had a huge response at the very beginning, but
Figure 11: Effect of the Frequency of Capacity Shocks

(a: Underweight, $p_u = 0.9$)  (b: Equal Weight, $p_u = 0.9$)

(c: Underweight, $p_d = 0.1$)  (d: Equal Weight, $p_d = 0.1$)

Figure 12: Effect of Frequency of Capacity Shocks with Equal Up and Down Probabilities

(a: Underweight)  (b: Equal Weight)
Table 8: Summary: Effect of Parameter Changes on Order Variance

| Increasing Element                                      | Impact on Order Variance |
|---------------------------------------------------------|--------------------------|
| Inventory level weight ($\alpha$)                        | Increase                 |
| Supply line weight ($\beta$)                            | First decrease, then increase |
| Target inventory level and supply line ($a, b$)         | Decrease                 |
| Exponential smoothing factor ($\eta$)                    | Increase                 |
| Stage-4 capacity ($\xi$)                                | Increase                 |
| Absolute value of reaction to capacity shock ($|\gamma_d|$) | Increase                 |

their purchasing behavior was smoother later, even though it took much longer for drilling and refining capacity to recover following the hurricane.

This difference between upstream and downstream can also be verified from Figure 12, where $p_u = p_d$ all the time; that is, the stationary probability of being up and down is the same. We set $\gamma = -10$, and $\xi_u$ and $\xi_d$ to 65 and 45 to ensure that the service rate is greater than the arrival rate. As $p_u$ and $p_d$ increase, disruptions become more frequent but shorter. As this happens, from Figure 12, we can see that the order SD at stage 1 and at stages 2 and 3 go in different directions. Even though the capacity variance is the same for different $p_d$ and $p_u$, in Figure 12, the inventory at stage 4 is less and less useful as the expected recovery time increases. That is why stages 2, 3, and 4 increase their order variance as $p_d$ and $p_u$ decrease. When $p_d$ and $p_u$ increase, disruptions occur more frequently, and therefore the order variance at stage 1 increases.

5.7 Discussion

In Table 8, we summarize the results of the previous sections by indicating the direction of change of the order variance as each element of the order function increases.

In our simulation study, we find that a relatively high weight on the supply line tends to result in stable performance throughout the supply chain. Therefore, if a company has a strong desire for production smoothing, then it is willing to keep its supply line as smooth as possible. If a company has a tight capacity limit, such as production capacities at its manufacturing sites or space limitation at its retailers, then it tends to absorb its clients’ order variability. The downside of this behavior is that the company’s clients may suffer insufficient supply, while its suppliers benefit from a smooth order pattern.

A company facing high holding costs may be more willing to apply just-in-time (JIT)
concepts, which means that it will reduce the target inventory level $\alpha$ and/or use a large weight $\alpha$ on the inventory level. Our study demonstrates that this will cause the company’s suppliers to suffer heavy BWE.

There is a substantial difference between a new product and a mature product. For a new product, a company may put more weight on the most recent realized demand in forecasting the future demand. In contrast, the company may be more inclined to use historical demand information in forecasting the demand of a mature product. Therefore, the BWE may be more severe for a new product than a mature product since we have demonstrated that BWE increases with the exponential smoothing coefficient $\eta$.

When there is a disruption, ordering responses may be different, depending on the importance of disrupted product to the customer. The more important the product is, the more likely the customer will be to over-order, while the less important the product is, the more likely the customer will be to under-order or simply keep the order size the same.

6. Conclusion

People drive any business. Therefore, behavioral supply chain research addresses the question of how people behave in various settings and the effect of that behavior on the supply chain as a whole. Recent research on the beer game experiment explores basic elements of human decision making, such as underweighting of the supply line. However, it is impossible to find a universal behavior that all managers follow.

Our studies use a beer game experiment to learn about the structure of human decision making, then perform simulation studies to examine the relationship between individual behavior and order variance quantitatively. Our descriptive models can serve as a foundation for studies involving questions of supply chain design and management. Just as in queueing theory, in which human behaviors of patience and abandonment can be quantified and used to develop optimal rules for designing queueing systems, we believe that future research should be performed to answer questions about the optimal form of ordering decisions when other managers in the system behave in a certain way.

In a dynamic world, managers face uncertainty not only from the demand side but also from the supply side. The past several years have seen a range of high-profile disruptions or near-disruptions, including Y2K, September 11th, SARS, the Indian Ocean tsunami, and hurricane Katrina. These low-probability, high-impact events have a tremendous impact
on the supply chain, as do smaller, less newsworthy disruptions that happen on a regular basis. In this paper, we studied potential forms of ordering behavior during disruptions by introducing supply uncertainty into the beer game and simulation experiments.

From our beer game experiment and simulation studies, we conclude that the BWE is not a ubiquitous phenomenon. We have identified three independent ways to generate the RBWE: over-weighting of the supply line, capacity constraints, and over-reactions to capacity shocks. The first two cause the RBWE by smoothing the order pattern upstream. The third generates an upstream shock that propagates downstream, and it may occur whether stages tend to under- or over-order during disruptions. All three of these causes provide some explanation as to why recent empirical studies have concluded that the BWE is not as prevalent as previously thought [2, 5]. From our simulation study, we also conclude that higher target safety stock levels and smaller exponential smoothing factors can effectively decrease the magnitude of the BWE. In addition, the pattern of orders in a supply chain appears to be more volatile when the behavior of the downstream stages is more irrational.
References

[1] Impact of hurricane katrina: One month after. Technical report, Information Resources, Inc., 2005. http://www.gmabrands.com/publications/gmairi/2005/special/ katrinaspesi-
cial4.pdf.
[2] Manuel P. Baganha and Morris A. Cohen. The stabilizing effect of inventory in supply chains. *Operations Research*, 46(3S):72–83, 1998.
[3] Elliot Bendoly, Karen Donohue, and Kenneth L. Schultz. Behavior in operations management: Assessing recent findings and revisiting old assumptions. 2006.
[4] Gary E. Bolton and Elena Katok. Learning-by-doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. 2005.
[5] Gérard P. Cachon, Taylor Randall, and Glen M. Schmidt. In search of the bullwhip effect. 2005.
[6] Dean C. Chatfield, Jeon G. Kim, Terry P. Harrison, and Jack C. Hayya. The bullwhip effect – impact of stochastic lead time, information quality, and information quality and information sharing: A simulation study. *Production and Operations Management*, 13(4):340–353, 2004.
[7] Fangruo Chen and Rungson Samroengraja. The stationary beer game. *Production and Operations Management*, 9(1):19–30, 2000.
[8] Frank Chen, Zvi Drezner, Jennifer K. Ryan, and David Simchi-Levi. Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management Science*, 46(3):436–443, 2000.
[9] Frank Chen, Jennifer K. Ryan, and David Simchi-Levi. The impact of exponential smoothing forecasts on the bullwhip effect. *Naval Research Logistics*, 47:269–286, 2000.
[10] Rachel Croson and Karen Donohue. Experimental economics and supply-chain management. *Interfaces*, 32(5):74–82, 2002.
[11] Rachel Croson and Karen Donohue. Impact of pos data sharing on supply chain management: An experimental study. *Production and Operations Management*, 12(1):1–11, 2003.
[12] Rachel Croson and Karen Donohue. Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science*, 52(3):323–336, 2006.

[13] Rachel Croson, Karen Donohue, Elena Katok, and John Sterman. Order stability in supply chains: Coordination risk and the role of coordination stock. 2004.

[14] Russell Gold, Bhushan Bahree, and Thaddeus Herrick. Storm leaves gulf coast devastated; rising oil and gas prices add to energy pressure on broader economy; ripples of a supply-side shock. *Wall Street Journal*, August 31 2005.

[15] P. Kaminsky and David Simchi-Levi. A new computerized beer game: A tool for teaching the value of integrated supply chain management. pages 216–225, 2000.

[16] Claudia Keser and Giuseppe Paleologo. Experimental investigating of supplier-retailer contract: The wholesaler price contract. 2004.

[17] Hua L. Lee, V. Padmanabhan, and Seungjin Whang. Comments on ”information distortion in a supply chain: The bullwhip effect”. *Management Science*, 50(12):1887–1893, 1997.

[18] Hua L. Lee, V. Padmanabhan, and Seungjin Whang. Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43(4):546–558, 1997.

[19] J. Mouawad. Energy producers make case for more coastal drilling. *New York Times*, October 14, 2005:C1, 2005.

[20] Jay Mouawad and Simon Romero. Gas prices surge as supply drops. *New York Times*, page A1, September 1 2005.

[21] Rogelio Oliva and Paulo Gonzalves. Evaluating overreaction to backlog as a behavioral cause of the bullwhip effect. 2006.

[22] Maurice E. Schweitzer and Gérard P. Cachon. Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science*, 46(3):404–420, 2000.

[23] Lawrence V. Snyder. BaseStockSim software v2.4. www.lehigh.edu/ lvs2/software.html, 2006.
[24] Joel H. Steckel, Sunil Gupta, and Anirvan Banerji. Supply chain decision making: Will shorter cycle times and shared point-of-sale information necessarily help? *Management Science*, 50(4):458–464, 2004.

[25] John D. Sterman. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35(3):321–339, 1989.

[26] Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131, 1979.

[27] Diana(Yan) Wu and Elena Katok. Learning, communication, and the bullwhip effect. 2005.