Behavioral Causes of Demand Amplification in Supply Chains: "Satisficing" Policies with Limited Information Cues

Rogelio Oliva
Mays Business School
Texas A&M University
Wehner 301F – 4217 TAMU
College Station, TX 77843
Tel (979) 862-3744; Fax (979) 845-5653
roliva@tamu.edu

Paulo Gonçalves
School of Business Administration
University of Miami
404 Kosar-Epstein Bldg.
Coral Gables, FL 33124
Tel (305) 284-8613; Fax (305) 284-2321
paulog@miami.edu

Draft: April 14, 2005.

Behavioral Causes of Demand Amplification in Supply Chains: "Satisficing" Policies with Limited Information Cues

Overreaction to supply shortages can create havoc in supply chains, costing millions of dollars in excess inventory and manufacturing capacity. In an experiment with the Beer Distribution Game, we explore overreaction to shortages as a complementary behavioral cause of supply chain instability. As in previous studies, we find that players ignore the supply line. We find, however, that instead of overreacting to shortages, players limit the size of their order adjustment while aiming for higher than necessary inventory level; a policy that is more stable than the linear response suggested in previous studies.

Since an ordering rule that fails to account for the supply line leads to higher than necessary costs and order amplification, our results suggest that players are not fully rational. However, evaluating the performance of the estimated policy we find that, given the information cues available, players show bounded rationality and develop a "satisficing" replenishment decision rule that minimizes local cost at the expense of higher upstream cost. We explore the implications of these findings for the design of information and incentive systems for supply chain management.

1. Introduction

Supply shortages are common in supply chains, often taking place in industries characterized by costly capacity and long acquisition delays (Cachon and Lariviere 1999), and during the introduction of new products, when demand is uncertain, and new processes, when production yield is uncertain (Lee et al. 1997a). More important, shortages can lead retailers to order in excess of their needs (Armony and Plambeck 2003; Gonçalves 2003; Lee et al. 1997a; Sterman 2000). The essence of over-ordering was captured by Mitchell (1924, p. 645) early last century:

> [R]etailers find that there is a shortage of merchandise at their sources of supply. Manufacturers inform them that it is with regret that they are able to fill their orders only to the extent of 80 percent. … Next season, if [retailers] want 90 units of an article, they order 100, so as to be sure, each, of getting the 90 in the pro rata share delivered."

As Mitchell suggests, when competing with other retailers for scarce supplies (i.e., horizontal competition), retailers inflate their orders to manufacturers seeking to improve their chances of obtaining the supply they need. Since manufacturers often sell to several retailers and shortages are common, overordering is a recurring problem in supply chains. Consider the following examples. Orders for DRAM chips skyrocketed after product shortages during the 1980s (Li 1992). Apple Computer regularly experienced inflated orders that were cancelled as soon as scarce products became available (Schneidawin and Cauley 1993) and Hewlett-Packard lost millions of dollars in unnecessary capacity and excess inventory following a post-shortage demand surge for its LaserJet printers (Lee et al. 1997b). More recently, Cisco Systems incurred a more than US$ 2 billion inventory write-off due to inflated retailer orders for its products (Adelman 2001).

Behavioral decision theory provides clues about how commonly used heuristics (e.g., representativeness and availability) might prompt overreaction (Kahneman et al. 1982). Overreaction typically results from excessive extrapolation of perceived patterns in random sequences or occurs in response to a dramatic or vivid event. For instance, De Bondt and Thaler (1985; 1987) found evidence supporting the overreaction hypothesis in financial markets. Frieder (2003) showed that individuals extrapolate past strings of positive news and trade too aggressively too late in reaction to extreme earnings announcements. And Massey and Wu (2004a; 2004b) found that people overreact when faced with imprecise information about a slowly changing environment.

In operations management, empirical studies have sought to capture overreaction in supply chains to explain the sources of the bullwhip effect*:* order amplification at each successive echelon upstream. Sterman's (1989) investigation of managerial decisions using the Beer Distribution Game (BDG) revealed that subjects failure to fully account for orders-placed-but-not-yet-received caused overordering. Several subsequent empirical studies have used the BDG to study factors influencing supply chain instability (Croson and Donohue 2002; 2003; Croson et al. 2004). A common finding across many of these studies has been the source of supply chain instability: players underestimate the supply line of orders placed but not received. This paper contributes to the empirical research on behavioral decision making within supply chains by articulating a complementary behavioral source of supply chain instability: overreaction in response to shortages.

Despite the lack of retailer competition for scarce supplies in the BDG, but motivated by Tversky and Kahneman's (1974) availability heuristic (i.e., the tendency to overreact to dramatic or vivid events) we hypothesize that players over-order when a shortage occurs because a backlog (a) is more costly than holding inventory, and (b) causes great disruption to the supply chain. Backlog is salient in both the cost objective function and the week-to-week operation of the supply chain. BDG facilitators emphasize the cumulative nature and financial impact of backlogs, including disruptions and stress to downstream players in need of stock. Finally, typical player quotes like “You better have something big on the way, or else!” and “Send anything! Please! We are thirsty” and ample testimonials during game debriefings (Sterman 1989) show that players become highly frustrated with the lack of supply, further increasing the appeal of the overreaction hypothesis.

We structure the data from 25 beer distribution games as a cross-sectional time-series panel, allowing us to make estimations across individuals and echelons. Our estimated ordering rule provides stronger evidence than previous studies that players underestimate the supply line. In fact, we find no significant effect of the supply line in the decision rule for non-factory echelons. Contrary to our expectations, we find that players do not overreact when in backlog, instead their correction saturates at a maximum value; a policy more stable than the linear response to shortfall suggested in previous studies. Through simulations we find that the estimated policy does not differ in form and cost performance from the local-cost-minimizing policy with the information cues used by players. Players, however, aim for a higher desired inventory than the local-cost-minimizing-policy, resulting in order amplification. Given the information cues available, we find that players show bounded rationality and develop a “satisficing” replenishment decision rule that minimizes their local cost at the expense of amplifying the demand signal. These findings have implications for the design of information and incentive systems in supply chain management. From a theoretical perspective, the paper contributes to the research on behavioral causes of the bull-whip effect. Through our analysis we identify two behavioral causes of demand amplification complementary to underestimating the supply line: higher than necessary desired inventory and overreaction to backlog. Players’ preference for the former is consistent with the low backlog-to-inventory cost ratio used in the BDG.

The remainder of the paper is structured as follows. In the section that follows we explore the empirical research on the causes of supply chain instability and the challenges of decision making in dynamic environments. In §3 we present our experimental design, in §4 our methods and results. In §5 we perform sensitivity analysis on the estimated stock management policy, compare its performance to the local-cost-minimizing policy, and evaluate the impact of backlog-to-inventory cost ratio on the cost-minimizing policy and our findings. We conclude with a summary of our findings and exploration of the implications for supply chain management and behavioral decision theory.

2. Dynamic Decision Making in Supply Chains

Decision making in supply chain management is highly complex and dynamic. To deal with such complexity, analytical models traditionally assume that rational agents optimize a well-defined and commonly-known utility function. If these assumptions are violated, however, model prescriptions might fail. As several empirical studies suggest, failure to meet the prescriptions of rational models is often the norm in a wide variety of fields such as economics (Kahneman et al. 1982; Plott 1986; Smith 1986), finance (Kahneman et al. 1986; 1990; Shiller 1981; Thaler 1988), marketing (Glazer et al. 1992) and operations management (Diehl and Sterman 1995; Schweitzer and Cachon 2000; Sterman 1989). Behavior remains sub-optimal even in dynamic environments, where the decision maker has supposedly the opportunity to identify and correct errors (Hogarth 1981). Experimental research suggests that decision makers perform poorly in environments with significant feedback delays (Sterman 1987; 1989), feedback complexity (Diehl and Sterman 1995; Schweitzer and Cachon 2000; Sterman 1989), and changing conditions (Kleinmuntz and Thomas 1987). Kleinmuntz (1993) concludes that decision makers are often "insensitive to the implications of feedback in dynamic environments."

Due to the dynamic nature of decision making in supply chain management, experimental research has the potential to address managers' cognitive limitations, providing a descriptive theory of how managers behave while guiding performance improvement. Moreover, current discrepancies between descriptive and normative research findings addressing supply chain instability calls for additional research. Consider the following examples. From a descriptive perspective, Sterman's (1989) experimental study of decision making in the Beer Distribution Game suggested that agents' inability to fully account for the supply line of orders placed-but-not-yet-received was responsible for supply chain instability. From the normative perspective, Lee et al. (1997a) highlight four operational causes for demand amplification generated in supply chains by rational agents: demand signal processing; rationing (supply shortages); order batching; and price variations. In support of the demand signal processing explanation, Chen et al. (2000) showed that in the presence of order lead times a simple forecasting rule could lead to demand amplification. In a setting similar to that of the Beer Distribution Game, Chen (1999) demonstrated that a base stock policy whereby managers place orders equal to those they receive minimizes total supply chain cost and avoids demand amplification when the demand distribution is stationary and commonly-know. Further experimental research found, however, that demand amplification persisted in an idealized supply chain even after controlling for all four operational causes suggested by Lee et al. (1997a) (e.g., Croson and Donohue 2002; 2003), and occurs even when demand is *fixed*, commonly-known, and players start at the *optimal* inventory level (Croson et al. 2004).

While the operational causes are important and clearly influence supply chain instability, even after controlling for them a number of behavioral reasons remain making the ordering decision in this dynamically complex environment far from trivial. For instance, it may be difficult to incorporate an information cue that is not readily available to decision makers, despite its importance to the task. Moreover, decision makers may generate only an incomplete set of possible decision rules or have faulty mental models about the environment (Kleinmuntz 1993); they may question if the optimal decision rule is commonly known and effectively used by other participants (Croson et al. 2004); and they may overreact in response to a dramatic or vivid event, as in the case of a sufficiently large order. This paper informs the research on causes of demand variability by articulating and analyzing a complementary behavioral source of instability: overreaction in response to shortages.

3. Experimental Design

Our experiment utilizes a web-based version of the Beer Distribution Game developed at Harvard Business School that maintains the essential structure of the board game (Sterman 1989). The game represents a serial supply chain with four echelons: retailer, wholesaler, distributor, and factory (R, W, D, and F, respectively). Each supply chain is independent of the other (echelons face no horizontal competition) and managed by a team charged with minimizing the supply chain cost. Each echelon incurs an inventory holding cost of $0.50 per unit/week and a backlog cost of $1.00 per unit/week. Shipment and order delays between echelons are two weeks and factories incur a one-week production delay with no capacity constraints. Players are randomly assigned, in pairs, to echelons (R, W, D, or F) and teams. Each simulated week players face the following sequence of events: (1) receive shipments; (2) fill customer orders, if sufficient inventory is available, otherwise accumulate a backlog; and (3) place an order with its supplier. Team members interact via a computer screen and, in contrast to the board version of the game, lack both visual access to the state of the supply line and knowledge of who their teammates are.

The game is initialized in flow equilibrium: order and shipment flows are 4 units/week and each echelon starts with an initial inventory level of 12 units. Subjects are not informed about the shape of demand. A single time increase in retailer orders (a step input) is introduced in the second period (week), bringing orders to 8 units/week. To avoid end-of-horizon behavior the experiment is announced to run for a simulated year, but is, in fact, terminated after 36 weeks. The web-based version, by virtue of its automatic computation of order receipts, incoming orders, shipments, and inventory-backlog levels, can be run with less time pressure than the board version of the game on which a facilitator imposes the pace. The automatic recording of transactional data avoids reporting errors, although data entry (i.e., “typing”) errors are still possible.

Our data set consists of 29 games played by first-year MBA students at Harvard Business

School during the fall of 2003. The game was played as part of the introductory course in operations management. We eliminated four outlying games from our sample. Three games had order variances above three standard deviations from the rest of the sample and one had a player with anomalous ordering behavior (consistent ordering of large quantities).[i]

3.1 The stock management problem

Sterman (1989) provided a general framework for the dynamic decision-making task of regulating a stock or system state and showed how that framework mapped into tasks as diverse as managing capital investments and personal energy level. This framework assumes a two stock structure with corresponding flows and a decision rule to manage the stock levels. Sterman proposed a simple, self-correcting decision making rule that uses information locally available to the decision maker and presumes no knowledge of the structure of the system. Specifically, managers are assumed to size orders to (1) replace expected losses from stock, (2) reduce the discrepancy between desired and actual stock, and (3) maintain an adequate supply line of unfilled orders (see Figure 1). The decision rule is formalized as:

$$O_t = MAX(0, \hat{L}_t + \alpha_S (S^* - S_t) + \alpha_{SL}(SL^* - SL_t)) \tag{1}$$

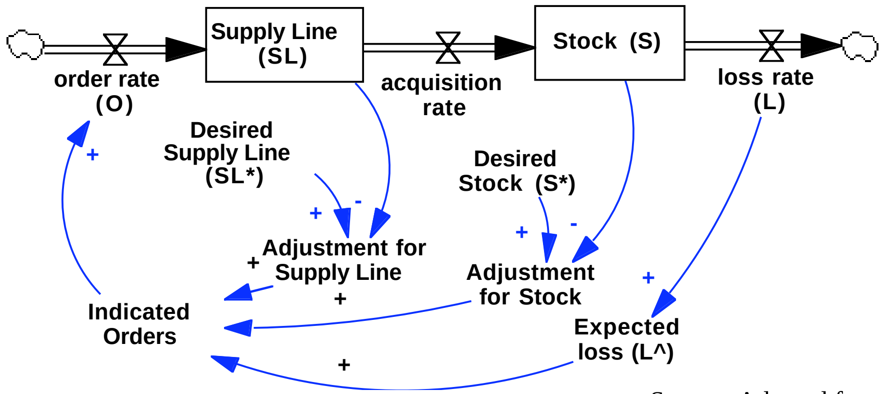
where, to be consistent with the BDG, orders are constrained to be nonnegative. $\hat{L}_t$ represents the expected loss from the stock, S_t and SL_t the inventory and supply line positions at time *t*, S^* and SL^* the desired levels for stock and supply line, and the parameters α_S and α_{SL} the fractional adjustment rate for inventory and supply line, respectively.

The remainder of this section explains how we revised the assumptions associated with each of the three components of Sterman's stock management decision rule to support the testing of our hypothesis.

Loss forecast. In his analysis of the decision rules used by BDG players, Sterman (1989) assumed adaptive expectations for the formation of the expected loss according to the exponential smoothing equation:

$$\hat{L}_t = \theta L_{t-1} + (1-\theta)\hat{L}_{t-1}. \tag{2}$$

Defining $\beta = \alpha_{SL}/\alpha_S$ and $S' = S^* + \beta SL^*$ Sterman obtained, for each player, maximum likelihood estimates for the simultaneous equations $O_t = MAX(0, \hat{L}_t + \alpha_S(S' - S_t - \beta SL_t) + \varepsilon_t)$ and $\hat{L}_t = \theta L_{t-1} + (1-\theta)\hat{L}_{t-1}$ subject to the constraints $0 \le \theta \le 1$ and $\alpha_S, S', \beta \ge 0$.



Source: Adapted from Sterman (1989).

Figure 1. Sterman's stock management problem

The joined estimation of these equations has the potential of shifting variance between the stock replenishment and forecasting equations, equations 1 and 2 respectively. Lower values of θ make the forecast series more stable and shift the residual variance to the replenishment decision (eq. 1), thus potentially biasing its parameter estimates (Oliva 2003). Since we did not have access to the forecasting rules used by BDG players, we made the simplifying assumption that players were only paying attention to the latest order they received (arguably the most salient one) and assumed a simple lag forecast $(\hat{L}_t = L_{t-1})$, an implied $\theta = 1$ in the exponential smoothing model in equation 2.

The simple lag forecast, an intuitive and plausible model of expectation formation (Kleinmuntz 1993), is consistent with the data and provided almost as good a fit as the optimal exponential smoothing. The simple lag assumption generated a reasonable forecast (0% Median Absolute Percent Error (MdAPE) and 23% Mean Absolute Percent Error (MAPE)) for a series with a coefficient of variation (σ/μ) of 0.61 and had a Root Mean Square Error (RMSE) only 9.5% higher than the optimal exponential smoothing forecast. [ii]

Inventory adjustment. Sterman (1989) treated backlog as negative inventory and assumed a linear response to the gap between current and desired inventory. Because the cost of backlog is twice the holding cost for inventory, it is possible that subjects reacted differently to backlog than to excess inventory. To test, with the simplest model possible, for the possibility of a different reaction to backlog, we assumed a piecewise linear model (Pindyck and Rubinfeld 1998), introducing a dummy variable (B_t) to reflect the backlog condition ($B_t = 1$ if $S_t < 0$; 0 otherwise). Accordingly, the response to inventory is modified to $\alpha_S\left(S^* - S_t\right) + \alpha_B S_t B_t$, where α_S represents the fractional adjustment rate for the inventory and α_B the incremental adjustment due to backlog (i.e., the response to backlog is $-\alpha_S + \alpha_B$). Since $S_t B_t \leq 0$, $\alpha_B < 0$ indicates a stronger reaction to the backlog condition (see Figure 2 for expected response to the inventory position and reaction to shortages).

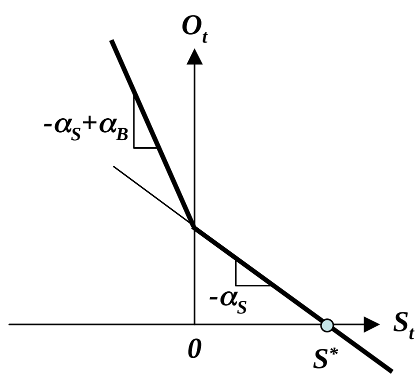


Figure 2. Order response to the inventory/backlog position

Supply line adjustment. Finally, Sterman assumed a negative relationship between orders and the supply line and, accordingly, constrained to negative values the search space for the fraction of the supply line taken into account (α_{SL}/α_S). He found this fraction to be significantly different from zero in 66% of his subjects and 0 to be the estimated value for 20% of the players. The average estimated value Sterman found for α_{SL} was -0.08 ($\alpha_{SL} = -\beta\alpha_S = -0.34 * 0.26$), but he did not report the significance of this estimate. He attributed the observed bullwhip effect in the BDG to players ignoring the information in the supply line.

We found a *positive* correlation between orders and supply line for the 100 players in our data set (r=0.26, p=0.000) that persisted when we split the sample by echelon, game, or player (see Table 1). The supply line being the cumulative difference between orders placed and shipments received, this positive correlation suggests an inverse causality from what Sterman estimated (i.e., that the supply line is large because large orders have been placed) and we decided to test its significance by omitting it from our base model of replenishment decision and later introducing it under various assumptions of goal formation.

Table 1. Correlation between orders and supply line, by echelon, game, and player

	Corr[O_t, SL_t]	Sig.		Corr[O_t, SL_t]>0	Corr[O_t, SL_t]<0
Retailers	0.24	0.000	Games (25)	25	0
Wholesalers	0.28	0.000	% Sig. at 0.05	76%	
Distributors	0.41	0.000	Players (100)	92	8
Factories	0.40	0.000	% Sig. at 0.05	53%	0%

Summarizing, our base model preserves the non-negativity constraint in orders but assumes a simple lag as the expected loss from stock, allows for a different inventory response when in backlog, and ignores the supply line. We formalized this base model for each player as:

$$O_t = MAX(0, L_{t-1} + \alpha_S (S^* - S_t) + \alpha_B S_t B_t) \tag{3}$$

where S_t, S^*, and α_S still represent the actual and desired inventory level and the fractional adjustment rate for the inventory, respectively; L_{t-1} is the expected loss; B_t is the backlog condition; and α_B represents incremental adjustment due to backlog. Defining $\beta_0 = \alpha_S S^*$, $\beta_1 = -\alpha_S$, and $\beta_2 = \alpha_B$, collecting terms, and allowing for an additive disturbance term yields a model with linear coefficients for inventory and the backlog condition:

$$O_t = MAX(0, L_{t-1} + \beta_0 + \beta_1 S_t + \beta_2 S_t B_t + \varepsilon_t). \tag{4}$$

4. Methods and Results

We treated the BDG's non-negativity constraint on orders as censored data. That is, we assumed that an order for zero could represent situations in which a subject wished to cancel a previously placed order (a negative order) but was restricted by the rules of the game to a minimum order of zero. Accordingly, we estimated the model using a tobit model (Tobin 1958). Finally, to estimate a decision rule that reflects the full range of observations available we structured the data from the games as a panel (cross-sectional time-series data set) with individual players the cross-sectional unit (*i*) and week of decision the time index (*t*). There being no reason to suspect that individual differences can be captured by changes in the constant term, and subjects being clearly a sample from a larger population, we assumed random effects across individuals (Greene

1997). With the expansion for panel data, our base model became:

$$O_{it} = MAX(0, L_{it-1} + \beta_0 + \beta_1 S_{it} + \beta_2 S_{it} B_{it} + u_i + \varepsilon_{it}) \tag{5}$$

where u_i is the random disturbance characterizing the ith subject. Estimations were performed using Stata's (2003) implementation of the random-effects cross-sectional time series tobit model and we tested the significance of the model's panel-level variance component (σ_u) by comparing the regression to the results of a pooled tobit regression.

4.1 Base model

Model I in Table 2 shows the estimated parameters for the base model together with the model's log-likelihood value, significance (χ^2), R^2, and root mean percent error. The model is highly significant and explains 60% of the variance in orders; 4% of that variance is explained by the differences among individuals $\left(\rho = \sigma_u^2 / \left(\sigma_u^2 + \sigma_\varepsilon^2\right)\right)$ and the panel-level variance is significant when compared to the pooled tobit model (likelihood-ratio test of $\sigma_u = 0$). Although we found some evidence of autocorrelation in the residuals that suggests that standard errors of estimates might be biased and the results not as efficient, tests on the linear model (i.e., without the non-negativity constraint) indicated that, for all regressions, the underestimation on the standard errors resulting from the autocorrelation was not large enough to affect the reported significance of the estimates—(in most cases the standard error of the estimates from the linear model matched that of the tobit model, and, when adjusting for autocorrelation in the residuals in the linear model, the change in the standard error of the estimates was in the third significant digit). Since estimates with autocorrelation are unbiased, the reported results can be safely interpreted.

The fractional adjustment rate of inventory and constant estimates have the expected sign; the estimated value for desired inventory ($S^* = \beta_0 / -\beta_1 = 7.94$) is consistent with one week of orders at the increased consumption rate; and the estimated fractional inventory adjustment ($\alpha_S = -\beta_1 = 0.21$) is consistent with, albeit a bit lower than, values found in previous studies that estimated versions of the stock management model for individual players ($\overline{\alpha}_S = 0.26$ in Sterman (1989) and $\overline{\alpha}_S = 0.23$ in Croson and Donohue (2002)). We were surprised, however, to find a positive coefficient for the backlog response ($\beta_2 = 0.21$). The combined response to a situation when inventory is in backlog ($\beta_1 + \beta_2$) is not significantly different from zero (see Backlog effect and its test in Table 2). This suggests that players in backlog place orders equal to the expected loss plus a constant amount (β_0) proportional the desired inventory level S^* (see Figure 3 for a schematic of the order response to the estimated model). Instead of “over-reacting” to a backlog situation as we had expected, players seem to ignore the backlog information cue and respond only to the $S_t = 0$ signal.

Table 2. Regression results

		Model I						Model II		
	Regressor	**Full**	**R**	**W**	**D**	**F**	**~F**	**Full**	**~F**	**F**
β_1	Inventory/Backlog	-0.21 (0.01)***	-0.14 (0.01)***	-0.18 (0.02)***	-0.20 (0.02)***	-0.40 (0.03)***	-0.17 (0.01)***	-0.21 (0.01)***	-0.15 (0.01)***	-0.43 (0.03)***
β_2	Backlog ($S_t B_t$)	0.21 (0.02)***	0.15 (0.03)***	0.22 (0.03)***	0.19 (0.03)***	0.34 (0.05)***	0.17 (0.01)***	0.21 (0.02)***	0.18 (0.02)***	0.31 (0.05)***
β_3	Supply line (SL_t)							0.00 (0.01)	0.02 (0.01)**	-0.09 (0.03)***
β_0	Constant	1.69 (0.17)***	1.30 (0.29)***	1.74 (0.28)***	1.60 (0.44)***	3.05 (0.57)***	1.44 (0.17)***	1.60 (0.25)***	0.77 (0.26)**	3.99 (0.67)***
	Log-likelihood value	-9426.0	-2161.7	-2367.5	-2455.3	-2128.1	-7193.6	-9425.8	-7188.4	-2123.3
	Wald χ^2	562.5***	110.7***	113.5***	139.5***	261.0***	349.8***	559.1***	348.2***	275.5***
	$R^{2(a)}$	0.60	0.16	0.43	0.49	0.77	0.43	0.60	0.43	0.77
	RMSE	4.46	2.94	3.88	5.33	5.11	4.18	4.46	4.17	5.09
	ρ	0.04 (0.01)**	0.07 (0.04)***	0.03 (0.02)	0.05 (0.02)***	0.10 (0.04)***	0.03 (0.01)***	0.04 (0.01)***	0.02 (0.01)***	0.12 (0.04)***
	Backlog effect P($\hat{\beta}_1 + \hat{\beta}_2 = 0$)	0.00 0.92	0.02 0.47	0.04 0.01	-0.01 0.45	-0.06 0.03	0.01 0.41	0.00 0.73	0.02 0.01	-0.12 0.00
	Observations	3500	875	875	875	875	2625	3500	2625	875
	Censored ($O_t \leq 0$)	520	30	76	154	260	260	520	260	260
	Number of players	100	25	25	25	25	75	100	75	25

Standard errors in parentheses: * significant at 10%; ** significant at 1%; *** significant at 0.1%.
(a) $R^2 = r^2$, where r is the simple correlation between estimated and actual orders (Wooldridge 2002).

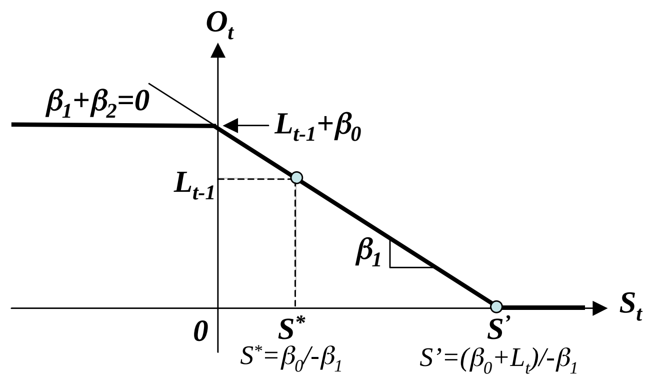


Figure 3. Estimated model's response to the inventory/backlog position

When the sample is split by echelon (see R, W, D, and F models in Table 2), the base model is significant for all positions and all the estimates have consistent signs and are significant (the model with dummies for each position yields the same results as Model I and the coefficients for the three dummies are not significant). The seemingly paradoxical result that both the R^2 and RMSE of the models increase as we move up the supply chain is explained by the fact that each successive stage the supply chain faces a demand stream with higher variance (see $\bar{\sigma}_{L_{t-1}}$ row in Table 3) and, although the models explain a higher fraction of that variance (R^2), the magnitude of the errors (RMSE) is increasing. In terms of parameter estimates, as a result of the differences in variance in demand each echelon faces, the aggressiveness of the inventory fractional adjustment ($\alpha_S = -\beta_1$) increases as we move up the supply chain: the higher the variance of L_{t-1}, the more aggressive corrections to deviations in inventory. Parameter estimates for the three first echelons (R, W and D), however, are not statistically different and the pooled model for these non-factory echelons (~F in Table 2) yields more efficient estimates, all within the standard error of the estimates for the separate models. Factories, whose stock management problem is structurally different from that of the other echelons—their delivery delay is shorter, and, because they are uncapacitated, it remains constant—have a significantly different replenishment rule with fractional adjustments to inventory more than twice as aggressive as the other echelons. The decision rule is, however, consistent across echelons as the other parameters adjust to accommodate the required aggressiveness of the fractional inventory adjustment—estimated values of S^* for each sample partition are not statistically different (see Table 3)—and maintain an almost-flat response to the backlog condition ($\beta_1 + \beta_2 \approx 0$).

Table 3. Demand variance and estimated parameters–base model and by echelon

	Full	R	W	D	F	~F
$\bar{L}_{t-1}$	**7.59**	**7.56**	**7.68**	**7.78**	**7.76**	**7.67**
$\bar{\sigma}_{L_{t-1}}$	4.78	1.26	3.23	5.11	7.30	3.56
$S^* = \beta_0 / -\beta_1$ [†]	**7.94**	**9.44**	**9.63**	**8.12**	**7.56**	**8.57**
S.E.	0.67	1.65	1.21	1.91	1.19	0.83

[†] Calculations based on the "delta method" (Oehlert 1992) values might differ from calculations based on coefficients from Table 2 because of rounding.

To test the robustness of this unexpected result we tested each of the assumptions made on our model specifications. First, we tested the forecasting assumption and replaced the simple-lag forecast with static expectations ($\hat{L}_{ti} = L^*$), extrapolative expectations ($\hat{L}_t = L_{t-1} + \gamma(L_{t-1} - L_{t-2})$, where $0 \le \gamma \le 1$), and optimal adaptive expectations ($\hat{L}_{it} = \theta_i L_{it-1} + (1-\theta_i)\hat{L}_{it-1}$, where θ_i minimizes the forecast error for each player). The almost-flat response to the backlog condition held under all forecasting assumptions. Second, we tested the non-linear shape of the response to inventory-backlog by testing different breakpoints for the piecewise linear model (the best fit was obtained when breakpoint is at S=0); introducing separate intercepts to the two line segments to decouple the inventory to the backlog response (second intercept was not significant), and testing different non-linear continuous responses (e.g., quadratic and logistic models)—all models generated a flat response to the backlog condition and none was as intuitive as the base model presented above. Third, we tested the censored data assumption and ran the model without the tobit constraint and found no significant change to the backlog response. Finally, we tested the aggregation assumption and ran the model, first, as a pooled data set (as expected from the small value of ρ, there were no significant change to the estimates), and then for each player independently: 84% of the players showed underreaction to backlog, and the model was significant for 67% of those players. The next section explores the impact of the supply line on the almost-flat response to the backlog condition.

4.2 Response to supply line

Model II in Table 2 shows the estimated parameters and performance statistics for a decision rule that incorporates the supply line signal:

$$O_{it} = MAX(0, L_{it-1} + \beta_0 + \beta_1 S_{it} + \beta_2 S_{it} B_{it} + \beta_3 SL_{it} + u_i + \varepsilon_{it}) \qquad (6)$$

For the full sample, the coefficient for the supply line is not significant and introducing it into the regression has no effect in the other estimates. When we split the sample in factories and non-factories, the *SL* coefficient becomes significant, taking the expected negative value for factories but a positive value for non-factories. Nevertheless, the introduction of the *SL* as a regressor does not have a significant impact in the inventory/backlog adjustment fractions for the two sub-samples, and the models' overall performance does not improve. When estimating model for individual players only 21% of the non-factories and 52% of the factories showed a significant negative response to the supply line. These results indicate that players consistently *ignore* the supply line, a finding stronger than Sterman's (1989) that players underestimate the supply line. We believe that Sterman's underestimation finding was obtained because he limited the search space for the *SL* coefficient to be negative. [iii]

4.3 Summary of findings

Our estimation strategy compares well to the strategy followed by previous studies with the BDG that estimate parameters for individual players (Croson and Donohue 2002; Croson et al. 2004; Sterman 1989). While Sterman's (1989) study yielded average R^2 and RMSE of 0.71 and 2.86 respectively, it required four parameters per player (i.e., 176 parameters for 44 players) to achieve this result. Our most parsimonious model achieves an R^2 of 0.60 and an RMSE of 4.46 with only three parameters for 100 players, confirming that it is possible to make some inferences across individuals. More importantly, the panel data structure integrated data across individuals and echelons, allowing us not only to obtain more efficient estimators but also to make formal inferences and tests across populations.

There are two interesting findings from this analysis. First, we found that players do not over-order when in backlog. Instead of "over-reacting" and having a more-than-proportional response to a backlog situation, players' adjustment saturates at a maximum, limiting the amount of amplification they introduce in the order stream. Second, our study confirms Sterman's finding that players underestimate the supply line. We find, however, that while factories are capable of keeping track of, and partially adjusting orders to, their supply line position (their response to the SL is similar to Sterman's average estimate), players in echelons with variable delivery delay have a *positive* response to the supply line. Since the inclusion of the supply line had no effect in the responses to inventory and backlog, and we did not find any evidence of endogenous adjustments to the desired inventory (S^*) and supply line (SL^*) positions, this positive response among non-factories suggests periods of slight over-ordering.

5. Sensitivity Analysis

A decision rule that does not consider the supply line amplifies orders and increases overall chain cost. Because orders placed to correct for an inventory imbalance do not arrive instantaneously, on-hand inventory remains low and correction orders are placed again and again resulting, eventually, in an overshoot of the inventory level. However, a decision rule that has a flat response to backlog limits the amount of amplification it introduces to the order stream. To understand the effects of this limited response, relative to the inherent amplification that results from ignoring the supply line, we explored the impact of different parameters of the players' decision heuristic on order amplification and cost.

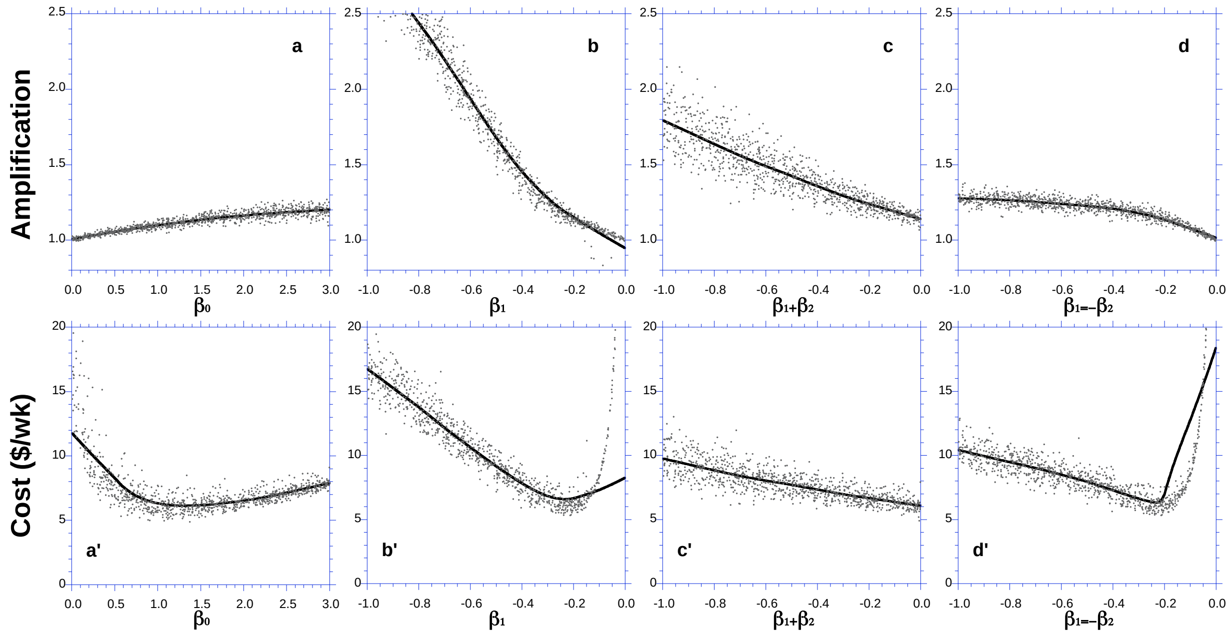
5.1 Sensitivity to decision parameters

To test the sensitivity of the different elements of the stock replenishment decision rule we used a simulated order stream. We used correlated noise with $O_t = MAX\left(0, \sim N\left(\mu, \sigma, \tau\right)\right)$ where $\mu = 8$ to reflect the steady state condition of the BDG; $\sigma = 4$ to approximate the incoming orders of our non-factory sample (see Table 3); and $\tau = 1.75$ the estimated autocorrelation time constant from orders received by non-retailers in our sample. We used two indicators to evaluate the performance of the decision rule. First, to assess the replenishment decision rule's contribution to the bullwhip effect we measured *order amplification* as the ratio of the standard deviation of the outgoing order stream (the orders placed by the decision rule) to the standard deviation of the incoming order stream. Second, to capture the inventory and service levels that resulted from applying the decision rule we used the cost structure stipulated in the rules of the BDG—$0.50 unit/week for holding inventory and $1.00 unit/week for backlog—to estimate the *average* *weekly cost*.

The simulated decision rule managed a position that faced, as in the BDG, a two-week information delay to communicate orders upstream and two-week transportation delay for orders to arrive. Because we assumed the supplier to have an infinite supply, any backlog incurred in this simulated environment is the result of structural delays and the decision rule. This assumption underestimates, relative to the BDG, the operating range for backlog for the simulated player. But because the decision rule does not take supply line into consideration and each decision made is based on the last period's order and current inventory/backlog position, the assumption does not affect the behavior of the decision rule per se and the results are comparable across simulations (tests with a model of the full supply line yielded qualitatively the same results).

Univariate sensitivity. To test the influence of each decision parameter we performed Monte-Carlo simulations, varying each parameter in isolation while maintaining the other parameters at the values estimated in the base case (Model I in Table 2). Each simulation was run for 300 weeks and since the model was initiated in equilibrium both measures of performance were calculated using data from the full simulation horizon. Figure 4 shows the resulting amplification and average weekly cost of 1,000 simulations varying each parameter in the displayed range.

The β_0 parameter represents the constant order quantity players would place in the absence of inventory and, with β_1, determines the desired inventory level ($S^* = \beta_0 / -\beta_1$). Being a constant term this parameter has a slight impact on order amplification (Figure 4a). Higher values of β_0 increase order amplification somewhat as the decision rule now has “more room” to adjust to excess inventory: with orders constrained to be positive the maximum inventory downsizing adjustment is limited to $-L_{t-1}$ when $S_t \geq (\beta_0 + L_{t-1}) / -\beta_1$ ($S_t \geq S^{'}$ in Figure 3); a larger β_0 increases $S^{'}$ and the operating range of the decision rule. As expected, the relationship between the target inventory level (S^*) and cost is u-shaped (Figure 4a’)—lower values of desired inventory result in a higher probability of running into a backlog situation and higher levels of inventory in excessive carrying cost.



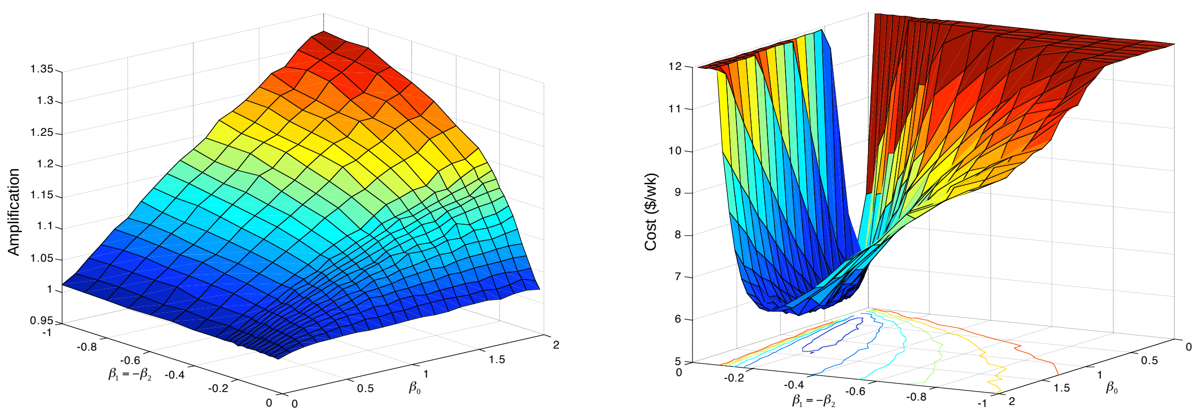
Smooth line estimated by means of the locally weighted regression scatter plot smoothing (lowess) procedure using, for each point, 40% of the sample (Chambers et al. 1983). The lowess line seems to deviate from the center of the data in the range (-0.21,0) in panels b and b' because the decision rule is unstable in that range and generated some extreme outliers. The deviation in panel d' is the result of extreme cost values when $\beta_1 \rightarrow 0$.

Figure 4. Univariate sensitivity of decision parameters—Monte-Carlo simulations

The fractional adjustment of the inventory position (α_S) is represented in the replenishment decision rule by $-\beta_1$. Note that the decision rule is not robust if the combined response to backlog ($\beta_1 + \beta_2$) is positive—if $\beta_1 + \beta_2 > 0$ the order response to backlog becomes weaker as backlog grows, eventually shutting down orders and allowing backlog to grow, resulting in ever increasing weekly cost (see Figure 3)—thus, the relevant range for β_1 is $(-\infty, -\beta_2]$. Figures 4b and 4b' show order amplification and average weekly cost decreasing in β_1 (i.e., as inventory correction becomes less aggressive)—more negative values of β_1 represent lower target inventory ($S^* = \beta_0 / -\beta_1$) increasing the probability of running into backlog.

Figures 4c and 4c' capture the combined value of $\beta_1 + \beta_2$ obtained by varying β_2 while maintaining β_1 at its estimated value. As in the case with inventory adjustment, amplification and cost are increasing in the intensity of the backlog adjustment. Although increasing the aggressiveness of the backlog adjustment increases the variability of the results, the average effect on performance is not as detrimental as when varying β_1. The minimum amplification and cost is obtained when $\beta_1 + \beta_2 = 0$, suggesting that for a decision rule that does not take supply line into consideration the optimal response to backlog is *no* response. Figures 4d and 4d' show the sensitivity of the decision rule when varying both parameters simultaneously ($\beta_1 = -\beta_2$). Variations in the response to inventory have little effect on amplification. The cost function, however, is highly sensitive to the effect that β_1 has on the implied target inventory ($S^* = \beta_0 / -\beta_1$).

Multivariate sensitivity. To separate the mutual dependency among parameters in the decision rule we performed a sensitivity analysis, varying both β_0 and β_1 while holding to the assumption of no backlog response ($\beta_2 = -\beta_1$). Figure 5 shows the surfaces of average amplification and average weekly cost as functions of β_0 and β_1 for 50 realizations of the incoming order stream. The general trends of the response surface are clear. Amplification is increasing in β_0 and in the intensity of the fractional adjustment of inventory, but there is a minimum cost basin at $\beta_0 \approx (0.5, 1.5)$ and $\beta_1 \approx (-0.1, -0.2)$, values consistent to the ones estimated from our non-factory sample.



Note that x and y axes are reversed in the cost graph to give a better perspective of the response surface and that isocost curves are shown in the xy plane. Cost values have been truncated at 12 to show better details of the minimal basin.

Figure 5. Multivariate sensitivity of decision parameters—Monte-Carlo simulations

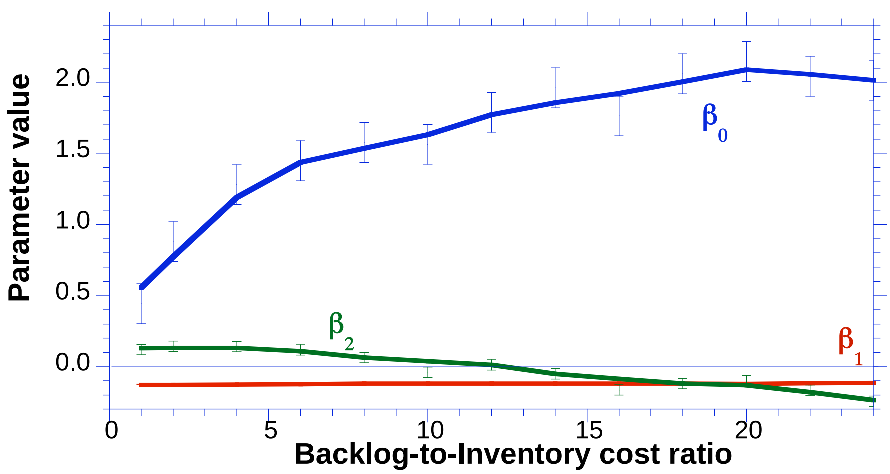
5.2 Comparison to local-cost-minimizing policy

To further assess the performance of the estimated policy we compared it to the local-cost-minimizing policy with the same information cues. For 50 different order realizations we found through a grid search of the parameter space the parameter values that would minimize cost. The average values for the cost-minimizing parameters β_0^*, β_1^*, and β_2^* were 0.89, -0.14, and 0.14, with standard errors of 0.06, 0.01, and 0.02, respectively. Confirming our interpretation of the sensitivity analysis, we found that for the local-cost-minimizing policy the response to backlog was not statistically different from 0 (p=0.339 for $H_0 : \beta_1^* + \beta_2^* = 0$). The response to the inventory position estimated from our non-factory sample ($-\beta_1 = \beta_2 = 0.17$) was slightly more aggressive than that of the local-cost-minimizing policy (p=0.003 for $H_0 : \beta_1 = \beta_1^*$, and p=0.067 for $H_0 : \beta_2 = \beta_2^*$), but the main difference between policies was that the estimated policy took a more conservative target inventory ($S_o^* = 6.36 < S_e^* = 8.47$) to reduce the probability of backlog (p=0.001 for $H_0 : \beta_0 = \beta_0^*$).

Evaluating the performance of the policies to 50 different order realizations revealed the average cost for the estimated rule to be only 2.5% higher than, the local-cost-minimizing policy, a non-significant difference (p=0.147 for $H_0 : c_o = c_e$). Order amplification, however, was found to be 4% higher for the estimated policy ($A_o = 1.07 < A_e = 1.12$), a highly significant difference (p=0.000 for $H_0 : A_o = A_e$). Note that the tests applied the estimated policy consistently throughout the simulation and that this consistency resulted in better performance than that of most players in our sample (Bowman 1963). Nevertheless, these results suggest that players were quite accurate in devising a rule that would minimize local cost given the information cues available. This cost minimization, however, came at the expense of increasing variance in upstream orders.

5.3 Sensitivity to backlog-to-inventory cost ratio

The results of the cost-minimizing policy suggest that a possible explanation for the flat response to backlog in our estimated rule is that while backlog is salient and twice as costly as inventory (i.e., a backlog-to-inventory cost ratio of 2) it is not costly enough to promote overreaction. Arguably real life costs associated with shortages (i.e., lost sales, lost goodwill, etc.) are much larger than twice the inventory holding costs. To test this, we explored the behavior of the local-cost-minimizing policy with the same information cues as the estimated rule through a range of cost ratios between 1 and 24. As above, we used a grid search to identify the parameter set that would minimize local cost. Figure 6 presents the general trend and standard errors for the cost-minimizing parameters β_0^*, β_1^*, and β_2^*. As the backlog-to-inventory cost ratio increases, the cost-minimizing rule adjusts by increasing the desired inventory level to reduce the probability of running into backlog; this is achieved by increasing β_0 since throughout the plotted range there is no sizable impact on the inventory response coefficient β_1. However, as desired inventory gets large enough to cover the order variance the benefit of increasing desired inventory decreases, resulting in a diminishing change rate for β_0. As the growth rate for β_0 decreases, β_2 begins to drop—suggesting more aggressive reaction to the backlog condition—and we see evidence of overreaction to backlog ($\beta_2 < 0$) for cost ratios larger than ten. The "late adjustment" of β_2 in Figure 6 supports the hypothesis that the backlog-to-inventory cost ratio used in the BDG is not sufficient to cause overreaction to backlog.



Smooth line estimated by means of the locally weighted regression scatter plot smoothing (lowess) procedure using, for each point, 35% of the sample (Chambers et al. 1983). Error bars represent the standard error of the parameter for a sample of 50 different order streams for each cost ratio.

Figure 6. Local-cost-minimizing policy's sensitivity to backlog-to-inventory cost ratio

6. Discussion

We explored, in an experimental serial supply chain, the causes of the bullwhip effect by proposing a complementary behavioral source of supply chain instability: overreaction to backlogs. The paper contains several contributions relative to previous work in this area. At the methodological level, refinements in assumptions and estimating techniques shed light on important aspects of the estimated decision rules for BDG players. First, by removing potential biases that could be introduced by the joined estimation of the adaptive expectations in forecast and the stock replenishment rule, and constraints in the feasible space for adjusting the supply line, we found even stronger evidence that players underestimate the supply line. Whereas previous work (Croson and Donohue 2002; Croson et al. 2004; Sterman 1989) suggested that the supply line was under-accounted for, we found no significant effect of the supply line in the decision rule for the non-factory echelons. Second, by structuring the data as a panel (cross-sectional time series), we not only could make use of all the data available for estimating the replenishment decision rule, thereby increasing the efficiency of estimates and the representativeness of the resulting rule, but also perform analyses by echelon in the simulated supply chain. We found that as players faced increasing order variance (e.g., upstream in the supply chain) the inventory adjustment fraction became more aggressive. The estimated decision rule was nevertheless robust across echelons, yielding similar values for the desired inventory and response to the backlog condition. Third, exploring a non-linear response to the inventory position revealed that, contrary to our expectations, players do not seem to overreact when in backlog; instead, they have a measured response to it, saturating order adjustment at a maximum value. This result held across echelons and under various assumptions for the non-linear response, the forecasting process, and the aggregation and estimation procedures.

The fact that players do not account for the supply line clearly indicates that they are not fully rational optimizers. An ordering rule that fails to account for the supply line will lead to order amplification. Because orders placed to correct for an inventory imbalance do not arrive instantaneously, on-hand inventory remains low and correction orders are placed again and again resulting, eventually, in an overshoot of the inventory level. Although the computation to estimate the supply line is simple (cumulative difference between orders placed and orders received), the form players use to track inventory status and orders placed does not provide the space, nor is time given when playing the game, to perform this computation. The supply line information is not salient in the game cues, nor does it play a role in the cost function for the individual players; thus, it is not surprising that players tend to ignore it (Plous 1993). Furthermore, players are not completely naïve considering that they do not over-order when in backlog. By not responding to the backlog condition, players create an ordering policy that is more stable than the linear response to inventory discrepancies. When evaluating the performance of the estimated rule, we found that players act as boundedly rational using the information available to them in a policy that is not significantly different in form and cost performance from the policy that minimizes local cost. The estimated rule, however, aims to maintain a higher inventory level than the cost-minimizing rule. Higher desired inventory increases order amplification as it creates “more room” for the decision policy to adjust inventory and increase the size of order adjustment. Higher order variance, in turn, increases costs for upstream echelons. Our findings suggest that boundedly rational players adopt a policy that “satisfices” local cost minimization at the expense of upstream order amplification and higher team costs. The estimated ordering policy indicates a strong behavioral component to supply chain instability, i.e., it ignores the supply line and underreacts to backlog while aiming for higher than necessary inventory level.

From a theoretical perspective, the paper contributes to the research on behavioral causes of the bull-whip effect. Our results suggest two additional components to demand amplification complementary to the underestimation of the supply line identified by Sterman (1989). First, higher desired inventory levels, when ignoring the supply line, increase the flexibility of the stock replenishment rule, allowing for larger corrections (i.e., higher order variability) in the order stream. Second, overreaction to backlog, i.e., placing orders beyond the proportional response to the inventory shortages, will clearly create demand amplification. In our sample we only found evidence of the higher-than-necessary desired inventory levels. This result, however, is consistent with the response of the local-cost-minimizing policy under the backlog-to-inventory cost ratio used in the BDG. That is, the backlog-to-inventory cost ratio used in the BDG was not high enough to trigger an overreaction to backlog in a cost-minimizing policy with the same information cues as the estimated rule. The impact of different backlog-to-inventory cost ratios on behavioral replenishment rules remains an area for future research.

Our empirical findings have implications for the design of information and incentive systems in supply chain management. Failure to account for the supply line raises the question of whether access to relevant information improves the likelihood of its adoption and results in a more adequate framing of the decision rule. Previous research suggests that access to the relevant information cues alone is insufficient to guarantee its utilization. In Sterman (1987), Kampmann (1992), and Diehl and Sterman (1995) supply line information is fully visible and just as salient as other important cues, but the supply line information is still severely underestimated. Richardson and Rohrbaugh (1990) found that clearly emphasizing important information cues, instead of just providing access to it, could improve the outcome of the decision rules. When they advised participants to “adjust for prior orders not yet filled” and provided the outstanding order amount, performance improved. Interestingly, when they provided a reasonable judgment policy without the proper motivation for it, players did not use it. While further research is required, decision making in dynamic environments seems to benefit from clearly emphasizing relevant information cues.

Analogously, although facilitators urge players to minimize total supply chain cost, local information and incentives shift players’ attention and the ordering rule accordingly. Formally, the game lacks a mechanism to bring global indicators (e.g., the costs of amplifying upstream orders) to the forefront and, by making them more salient, increase the likelihood that they will be incorporated on the decision rule. Here too, lack of access to amplification cost information and clear understanding of its role likely justifies its absence from the adopted decision rule. Emphasizing the impact and cost of order amplification to upstream players might be an effective way to make the relevant information cues more salient. One possibility may be to contrast estimated team costs achieved using different policies by estimating the cost impact caused by upstream order amplification. The individual savings achieved by attempting to minimize local costs may be greatly overshadowed by higher team costs imposed due to increased order amplification to upstream players. Although it is now standard practice in supply chain design to account for supply line information and to increase inventory visibility throughout the chain, managers still attempt to optimize locally despite the detrimental impact that their decisions may have on the whole supply chain, including themselves through numerous different feedback mechanisms. Currently, managers do not internalize the impact that their decisions may have on upstream players and indirectly on themselves. However, by considering the cost of upstream order amplification (or other relevant information cues in a different context), managers may be able to improve their policy decision and performance. Access to information alone may prove less useful than understanding its impact on supply chain coordination. A promising area for further research is the role that emphasizing information cues or adequate motivation of specific decision policies might have on decision making in dynamic environments.

Acklowledgments

We are grateful to John Sterman, Benito Flores, Alan Love, Arzu Onar and participants of the INFORMS, POMS, and System Dynamics Winter Conferences, for valuable feedback on earlier versions of this work, and to Lisa Noonan for her help running the games and capturing the data.

References

Adelman, D. 2001. First Loss Sets Gloomy Tone at CISCO. *Wall Street Journal*, May 9.

Armony, M., E. Plambeck. 2003. The Impact of Duplicate Orders on Demand Estimation and Capacity Investment. Working Paper, Stern School of Business, New York U., New York.

Bowman, E.H. 1963. Consistency and Optimality in Managerial Decision Making. *Manage. Sci.* **9** 310-321.

Cachon, G.P., M.A. Lariviere. 1999. Capacity Choice and Allocation: Strategic Behavior and Supply Chain Performance. *Manage. Sci.* **45**(8) 1091-1108.

Chambers, J.M., W.S. Cleveland, B. Kleiner, P.A. Tukey. 1983. *Graphical Methods for Data Analysis*. Wadsworth Int. Group, Belmont, CA.

Chatfield, C. 2001. *Time-Series Forecasting*. Champman & Hall, London.

Chen, F. 1999. Decentralized Supply Chains Subject to Information Delays. *Manage. Sci.* **45**(8) 1076-1090.

Chen, F., Z. Drezner, J.K. Ryan, D. Simchi-Levi. 2000. Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information. *Manage. Sci.* **46**(3) 436-443.

Croson, R., K. Donohue. 2002. Behavioral Causes of the Bullwhip Effect and the Observed Value of Inventory Information. Working Paper, Carlson School, U. of Minnesota, Minneapolis.

Croson, R., K. Donohue. 2003. Impact of POS Data Sharing on Supply Chain Management: An Experimental Study. *Prod. Oper. Manag.* **12**(1) 1-11.

Croson, R., K. Donohue, E. Katok, J.D. Sterman. 2004. Order Stability in Supply Chains: Coordination Risk and the Role of Coordination Stock. Working Paper, Smeal College of Business, Penn State, State College, PA.

De Bondt, W., R. Thaler. 1985. Does the Stock Market Overreact? *J. Financ.* **40**(3) 793-805.

De Bondt, W., R. Thaler. 1987. Further Evidence on Investor Overreaction and Stock Market Seasonality. *J. Financ.* **42**(3) 557-581.

Diehl, E., J.D. Sterman. 1995. Effects of Feedback Complexity on Dynamic Decision Making. *Organ. Behav. Hum. Dec.* **62**(2) 198-215.

Frieder, L. 2003. Evidence on Behavioral Biases in Trading Activity. Working Paper, Anderson School of Business, UCLA, Los Angeles.

Glazer, R., J.H. Steckel, R.S. Winer. 1992. Locally Rational Decision Making: The Distracting Effect of Information on Managerial Performance. *Manage. Sci.* **38**(2) 212-226.

Gonçalves, P. 2003. Phantom Orders and Demand Bubbles in Supply Chains. Ph.D. Dissertation, Sloan School of Management, MIT, Cambridge, MA.

Greene, W.H. 1997. *Econometric Analysis*. 3rd ed. Prentice Hall, Upper Saddle River, NJ.

Hogarth, R.M. 1981. Beyond Discrete Biases: Functional and Dysfunctional Aspects of Judgments and Heuristics. *Psychol. Bul.* **90** 197-217.

Kahneman, D., J.L. Knetsch, R.H. Thaler. 1986. Fairness and the Assumptions of Economics. *J. Bus.* **59**(4) S285-S300.

Kahneman, D., J.L. Knetsch, R.H. Thaler. 1990. Experimental Tests of the Endowment Effect and the Coase Theorem. *J. Polit. Econ.* **98**(6) 1325-1348.

Kahneman, D., P. Slovic, A. Tversky. 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge.

Kampmann, C.E. 1992. Feedback COmplexity and Market Adjustment: An Experimental Approach. PhD Dissertation, Sloan School of Management, MIT, Cambridge, MA.

Kleinmuntz, D. 1993. Misperceptions of the Implications of Feedback in Dynamic Decision Making. *Syst. Dyn. Rev.* **9**(3) 223-237.

Kleinmuntz, D., J. Thomas. 1987. The Value of Action and Inference in Dynamic Decision-

Making. *Organ. Behav. Hum. Dec.* **39**(3) 341-364.
Lee, H., V. Padmanabhan, W. Seungjin. 1997a. Information Distortion in a Supply Chain: The Bullwhip Effect. *Manage. Sci.* **43**(4) 546-558.
Lee, H., V. Padmanabhan, S. Whang. 1997b. The Bullwhip Effect in Supply Chains. *Sloan Manage. Rev.* **38**(3) 93-102.
Li, L. 1992. The Role of Inventory in Delivery-Time Competition. *Manage. Sci*. **38**(2) 182-197.
Massey, C., G. Wu. 2004a. Detecting Regime Shifts: The Psychology of Under- and Over-Reaction. Working Paper, Fuqua School of Business, Duke U., Durham, NC.
Massey, C., G. Wu. 2004b. Understanding Under- and Over-Reaction. Working Paper, Fuqua School of Business, Duke U., Durham, NC.
Mitchell, T.W. 1924. Competitive Illusion as a Cause of Business Cycles. *Q. J. Econ.* **38**(4) 631-652.
Oehlert, G.W. 1992. A Note on the Delta Method. *Am. Stat.* **46** 27-29.
Oliva, R. 2003. Model Calibration as a Testing Strategy for System Dynamics Models. *Eur. J. Oper. Res.* **151**(3) 552-568.
Pindyck, R.S., D.L. Rubinfeld. 1998. *Econometric Models and Economic Forecast*. 4th ed. Irwin, McGraw-Hill, Boston.
Plott, C.R. 1986. Laboratory Experiments in Economics: The Implications of Posted-Price Institutions. *Science* **232**(4751) 732-738.
Plous, S. 1993. *The Psychology of Judgment and Decision Making*. McGraw-Hill, New York.
Richardson, G.P., J.W. Rohrbaugh. 1990. Decision Making in Dynamic Environments: Exploring Judgments in a System Dynamics Model-Based Game. In Brocherding, K., O. I. Larchev *et al.* (eds.), *Contemporary Issues in Decision Making*, Elsevier, Amsterdam, pp. 463-472.
Schneidawin, J., L. Cauley. 1993. Apple's New CEO Faces Hard Choices. *USA Today*, June 21.
Schweitzer, M.E., G.P. Cachon. 2000. Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. *Manage. Sci.* **46**(3) 404-420.
Shiller, R.J. 1981. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *Am. Econ. Rev.* **71**(3) 421-436.
Smith, V.L. 1986. Experimental Methods in the Political Economy of Exchange. *Science* **234**(4773) 167-173.
Stata. 2003. *Stata Release 8: Cross-Sectional Time-Series*. Stata Press, College Station, TX.
Sterman, J.D. 1987. Testing Behavioral Simulation Models by Direct Experiment. *Manage. Sci.* **33**(12) 1572-1592.
Sterman, J.D. 1989. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Manage. Sci.* **35**(3) 321-339.

Sterman, J.D. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill, New York.

Thaler, R.H. 1988. Anomalies: The Winner's Curse. *J. Econ. Perspect.* **2**(1) 191-202.

Tobin, J. 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica* **26** 24-36.

Tversky, A., D. Kahneman. 1974. Judgment Under Uncertainty: Heuristics and Biases. *Science* **185**(4157) 1124-1131.

Wooldridge, J.M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

---

[i] In this web-based version of the game orders are limited to the [0,99] range but players are not aware of the upper-limit unless they attempt to place an order greater than 99, in which case the software truncates the order and informs the players about it. Not revealing the upper limit ensures that players do not self-constraint the size of their orders based on this additional information cue. All orders in the 25 games in our sample were for less than 99 units.

[ii] We tested the significance of the exponential smoothing as a predictor using the result that the moving average (MA) coefficient (ω) of an ARIMA (0,1,1) model is equivalent to the smoothing parameter in a single exponential smoothing $(\theta = 1 + \omega)$ (Chatfield 2001). Applying the ARIMA (0,1,1) model to the loss series, we found the MA coefficient to be significant for only 14% of the players. That is, only for 14% of the players was θ found to be different from one. The optimal exponential smoothing forecast was determined by finding the value of θ in equation (2) that would minimize the forecasting error for each player.

[iii] A potential explanation for the positive and significant coefficient of the supply line for non-factories is that it reflects the net effect from a decrease in orders due to a large supply line (the expected negative coefficient) and an increase in orders due to a higher desired supply line (SL^*). Assuming an endogenous goal formation as a function of the loss forecast ($S_t^* = \kappa_S + \gamma_S \hat{L}_t$ and $SL_t^* = \kappa_{SL} + \gamma_{SL} \hat{L}_t$) in Sterman's original equation (eq. 1); replacing $\beta_0 = \alpha_S \kappa_S + \alpha_{SL} \kappa_{SL}$, $\beta_1 = -\alpha_S$, $\beta_3 = -\alpha_{SL}$ and $\beta_4 = 1 + \alpha_S \gamma_S + \alpha_{SL} \gamma_{SL}$; and reintroducing the backlog, panel and random variation terms, yields the linear model $O_{it} = MAX(0, \beta_0 + \beta_1 S_{it} + \beta_2 S_{it} B_{it} + \beta_3 SL_{it} + \beta_4 \hat{L}_{it} + u_i + \varepsilon_{it})$. Under the endogenous goal formation hypothesis $\alpha_S, \gamma_S, \alpha_{SL}$, and $\gamma_{SL} \geq 0$, thus we would expect $\beta_4 \geq 1$. We rejected this hypothesis since this model for the non-factories in our sample yielded $\beta_4 = 0.77$ (S.E.=0.03), i.e., p=0.999 for H_0: $\beta_4 < 1$.