MAIN ARTICLE

The impact of information visibility on ordering dynamics in a supply chain: a behavioral perspective

Paulo Gonçalves[a] and Mohammad Moshtari[b]*

Abstract

Previous research on the Bullwhip Effect shows that information visibility—Point-Of-Sale (POS) data or supply-chain partner-inventory data—can reduce the amplification of orders in a supply chain. This study compiles and analyzes the data from two previous experiments with the beer game (Croson and Donohue, 2003, 2006) to gain insight on the specific mechanisms that decrease order amplification. By structuring the data as a panel and using a fixed-effects estimation model, we find that additional supply-chain-level information (e.g., POS and inventory data) leads subjects to react less aggressively to changes in their own inventory and to pay more attention to the supply line of orders placed. Furthermore, our analysis shows that such effects are more pronounced for upstream subjects. Our findings provide insight into the role that additional supply-chain information play on subjects' orders in the beer game.

Syst. Dyn. Rev. **37**, 126–154 (2021)

Additional Supporting Information may be found online in the supporting information tab for this article.

Introduction

The Bullwhip Effect, the increase in order amplification as one moves upstream in a supply chain (Lee *et al.*, 1997) has attracted the attention of practitioners and academics for decades, motivating a number of studies with a wide variety of approaches and ultimate goals (e.g. Sethuraman and Tirupati, 2005; Croson and Donohue, 2006; Macdonald *et al.*, 2013; Khan *et al.*, 2019).

Prior research suggests two categories of causes triggering the Bullwhip Effect: operational and behavioral. Research focusing on operational causes of the Bullwhip Effect assumes that supply-chain actors are fully rational and optimize a single, agreed upon by all, objective. Lee *et al.* (1997) offer four operational causes for the Bullwhip Effect: price fluctuations, order batching, rationing game (due to supply shortages), and demand signal processing. Chen *et al.* (2000) quantify the Bullwhip Effect in a two-stage

[a] Institute of Management, Università della Svizzera italiana, Lugano, Switzerland
[b] Tampere University, Tampere, Finland
* Correspondence to: Mohammad Moshtari, Tampere University, Korkeakoulunkatu 8, FI-33101 Tampere, Finland. E-mail: mohammad.moshtari@tuni.fi

Accepted by Michael Shayne Gary, Received 23 October 2020; Revised 8 May 2021 and 7 July 2021; Accepted 19 July 2021

System Dynamics Review
System Dynamics Review vol 37, No 2-3 (April-September 2021): 126–154
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/sdr.1687

supply chain subjected to two of the causes: demand forecasting and order lead times. They also demonstrate that centralizing information can reduce, but not eliminate, the Bullwhip Effect. Other work in this operational area focuses on quantifying the Bullwhip Effect. Luong (2007) and Duc *et al.* (2008) quantify the Bullwhip Effect in a single-supplier single-retailer supply chain. The former assumes retailers use a base stock-inventory policy and an auto-regressive demand forecast. The latter assumes stochastic demand and lead times. Finally, Cachon *et al.* (2007) use industrial-level U.S. data, instead of a theoretical model, to measure and quantify the strength of the Bullwhip Effect in the economy.

In contrast to operational research where actors are fully rational, behavioral research assumes that supply-chain actors are boundedly rational, display limits on rationality and cognitive processing, and use heuristics to make decisions. Behavioral research finds that the Bullwhip Effect persists, even when researchers control for operational causes in their experiments. Behavioral causes of the Bullwhip Effect include underestimating the supply line (Sterman, 1989), reaction to backlog (Oliva and Gonçalves, 2007), phantom ordering (Sterman and Dogan, 2015), coordination risk (Croson *et al.*, 2014), and scope neglect (Oliva *et al.*, 2021). Different attributes can influence those behavioral causes such as the cognitive profile of people (Narayanan and Moritz, 2015), retailer shortages in a competitive environment (Villa *et al.*, 2015), and access to POS or inventory information (Croson and Donohue, 2003, 2005, 2006).

Bridging the potential divide between operational or behavioral causes of the Bullwhip Effect, Croson and Donohue (2003, 2005, 2006) observe that the Bullwhip Effect still exists in the beer game even in the absence of price fluctuations, order batching, and rationing game, i.e. three out of the four operational causes (Lee *et al.*, 1997).

Along this behavioral line of research, two studies by Croson and Donohue (2003, 2006) explored the influence of POS data and sharing inventory data, respectively, on the Bullwhip Effect. Their studies suggest that additional information (POS or inventory data) can significantly reduce order amplification and the Bullwhip Effect in the beer game, in particular among upstream echelons (Croson and Donohue, 2003, 2006). They posit that “access to the inventory levels of upstream members could improve a decision maker’s ability to anticipate supply shortages” possibly allowing him to “combat the supply line underweighting tendency” (Croson and Donohue, 2006, p. 330). Their results, however, show that supply line underweighting is still prevalent when additional information is available and that the differences in the supply-line coefficients, between the control and the treatment groups, were not significant. To explain this, they hypothesize that access to POS data can help “upstream suppliers to better anticipate their customers’ needs” (Croson and Donohue, 2003, p. 10) and, that “inventory information … improves performance by allowing manufacturers and

DOI: 10.1002/sdr

distributors to anticipate and interpret orders placed by their downstream customers" (Croson and Donohue, 2006, p. 333).

Hence, Croson and Donohue (2003, 2006) hypothesize, but do not test, mechanisms by which POS data and sharing inventory data reduce the Bullwhip Effect. The main goal of this study is to empirically explore specific mechanisms by which the availability of additional information in the beer game affects subjects' ordering behavior. We investigate the proposed mechanisms on subjects that are upstream and downstream in the supply chain. Methodologically, we use Croson and Donohue's (2003, 2006) experimental data and structure it as a panel (cross-sectional time-series data set) to estimate a single empirical model for ordering behavior that reflects all the available observations (from 128 subjects). The panel-data structure increases the efficiency of the estimates and the representativeness of the resulting rule as it allows us to make estimations across individuals and echelons. The findings with the panel structure are robust to individual estimation models.

Our research indicates two specific ways that additional information (POS or inventory data) affects ordering behavior. First, subjects are more conservative in eliminating inventory gaps, that is, they react less aggressively to their own inventory data. Second, subjects pay more attention to the supply line. Considering orders they receive, we find that additional information does not significantly affect subjects' use of forecasts to their own order policy. Moreover, our research shows that both effects are more pronounced for upstream subjects. Our research builds upon Croson and Donohue's (2003, 2006) findings to identify the specific mechanisms by which availability of additional information reduces the Bullwhip Effect.

Comparative overview

In this section, we provide a comparative overview of nine previous research articles investigating behavioral causes of the Bullwhip Effect using the beer game. Table 1 summarizes key aspects (e.g. sample size, method, assumptions, treatments, dependent variables, contributions) of those studies. As the table shows, the samples in the studies range from 10 to 60 games (40 to 240 inviduals). In terms of assumptions, studies typically assume that demand is either known and stationary or unknown and nonstationary. After the influential work of Lee *et al.* (1997) detailing operational causes of the beer game (e.g. price fluctuations, order batching, rationing game, and demand forecasting), behavioral operations studies control for as many operational causes as possible. Three studies without treatments, including Sterman's (1989) seminal work, attempt to make sense of the decision rule adopted by subjects and how they use available information cues (e.g. expected loss, inventory, backlog, supply line, etc.) in the game. Among

DOI: 10.1002/sdr

Table 1. Overview of recent beer game studies

Study	Sample[a]	Method	Assumptions	Treatment/modeling	Dependent variable	Key findings
Sterman (1989)	11 (4)	• Runs a set of regressions at individual (echelon) level	• Demand unknown and nonstationary • Controls three operational causes	No treatment	Order	• Bullwhip effect exits • "Underweighting to supply line" is observed
Croson and Donohue (2003)	11 (4) 10 (4)	• Runs a set of regressions at individual (echelon) level • Wilcoxon Test	• Demand known and stationary • Controls three operational causes	POS data sharing	Order variation	• The magnitude of order oscillations decreases • The magnitude of amplification of order decreases oscillation just between retailers and wholesalers • Upstream members experience more pronounced decrease in order oscillations • "Underweighting to supply line" is still observed • Less "overreaction to demand" is observed
Croson and Donohue (2005)	60 (4)	• Seigel Sign Test (Amplification Comparison) • Wilcoxon Test (order oscillations reduction)	• Demand known and stationary • Control three operational causes	Upstream versus downstream inventory information	Order variation	• Sharing downstream inventory information decreases order oscillations • There is more decrease in order oscillations for upstream members
Croson and Donohue (2006)	11 (4) 11 (4)	• Wilcoxon Test (compare across two treatments) • Runs a set of regressions at individual level	• Demand known and stationary • Controls three operational causes	Inventory information	Order variation	• Order oscillation decreases • Amplification of order oscillation does not decrease • "Underweighting to supply line" is still observed • There is more decrease in order oscillations for upstream members

(Continues)

DOI: 10.1002/sdr

Table 1. Continued

Study	Sample[a]	Method	Assumptions	Treatment/modeling	Dependent variable	Key findings
Oliva and Goncalves (2007)	25 (4)	• Tobit model—nonnegative orders as censored data • Structure data as a panel and use a fixed-effect • Runs a set of regressions at aggregate and echelon levels	• Demand nonstationary and unknown	No Treatment	Order	• Bullwhip effect exists • Supply-line underweighting • Under-reaction to backlog

Study	Sample	Method	Assumptions	Treatment/ Modeling	Dependent variable	Key Findings
Croson *et al.* (2014)	40(4)	• (Amplification Comparison) • Wilcoxon Test (order oscillations reduction)	• Demand constant and known	Four experiments) varying level of common knowledge of optimal order policy; automation of supply chain partner decisions, and initial level of on-hand inventory)	Order variation	• Coordination risk contributes to Bullwhip behavior • Performance is improved through the introduction of common knowledge, the addition of trust, and adding coordination stock, though underweighting the supply line still persists
Sterman and Dogan (2015)	60 (4) Croson *et al.* (2014) data plus two more experiments not reported in the study	• Runs set of regressions at the individual level	• Demand constant and known • Relax assumption of constant desired on-hand and on-order inventory levels	Six treatments)varying level of common knowledge of optimal order policy, automa-tion of supply chain partner decisions, and initial level of on-hand inventory)	Order	• Hoarding and phantom ordering particularly strong for the outliers who placed extremely large orders
Oliva, Abdulla, & Goncalves (2021)	25(4)	• Structures data as three level nested panel • Uses multilevel mixed-effect Tobit model	• Demand nonstationary and unknown • Demand stationary and known (robustness check)	No Treatment	Order	• Scope neglect - when in backlog subjects order less aggressively and become insensitive to problem scope

(Continues)

DOI: 10.1002/sdr

Table 1. Continued

Study	Sample	Method	Assumptions	Treatment/ Modeling	Dependent variable	Key Findings
This Study	11(4) 10(4) 11(4) Croson and Donohue (2003, 2006)	• Structures data as a panel and uses a fixed-effect model • Runs set of regressions at the aggregate and upstream/ downstream levels	• Demand stationary and known • Controls three operational causes	Inventory information and POS data sharing	Order	• Supply chain level information leads subjects to react less aggressively to own inventory data and to pay more attention to supply line • POS and inventory data cause upstream players to be less reactive to own inventory information and in inventory information treatment to pay more attention to the supply line of previous orders placed

aNumber of teams per experiment and number of players per team.

DOI: 10.1002/sdr

the different treatments, studies consider the impact of (a) POS data; (b) inventory information throughout the supply chain; (c) upstream and downstream inventory information; (d) common knowledge of optimal order policy; (e) automation of supply-chain partner decisions; and (f) initial level of on-hand inventory. A common result from treatments is that additional information, common knowledge, trust, and coordination stock reduce Bullwhip Effect and improve overall performance. Methodologically, several studies focus on individuals as the unit of analysis (e.g. Sterman, 1989; Croson and Donohue, 2003, 2005, 2006; Croson *et al.*, 2014; Sterman and Dogan, 2015), estimating coefficients for different information cues by computing averages across individuals in different positions. A few studies focus on the supply chain (e.g the beer game) as the unit of analyses (e.g. Oliva and Gonçalves, 2007, Oliva, Abdulla & Goncalves 2021, and this study), obtaining coefficients estimates for different information cues directly. The dependent variable across all studies focus on either the orders placed or the order variability.

Table A1 reveals that this study is well alligned with previous ones with respect to sample, methods, assumptions, treatments, and dependent variable. At the same time, this study provides novel insights regarding the role of information on reaction to own inventory and supply line.

Experimental setting

In this section, we review the experimental setting designed by Croson and Donohue (2003, 2006) to conduct beer game experiments with additional information. We do not conduct any new experiments in our study, using data available from Croson and Donohue (2003, 2006). Our use of data from previous research is similar to Sterman and Dogan (2015), who use data from Croson *et al.* (2014) to derive their findings. The beer-distribution game captures a serial supply chain with four echelons: retailer, wholesaler, distributor, and manufacturer (Figure 1). Each beer game team (i.e. supply chain team) consists of four subjects, who independently play one of four roles (i.e. retailer, wholesaler, distributor, or manufacturer). Each team seeks to minimize total supply-chain cost (i.e. the total cost for all four echelons for the duration of play). Appendix A.1, Appendix S1 in the online supporting information provide more information about experiment design.

Croson and Donohue (2003, 2006) consider three treatments:

- **Baseline treatment** (same treatment in Croson and Donohue (2003, 2006): Includes data for 11 teams (44 players) placing orders for 48 weeks. All subjects know the distribution of final customer demand. Only retailers observe final customer demand (POS data).

DOI: 10.1002/sdr

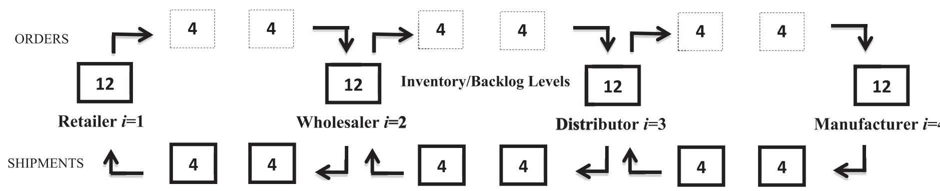

Fig. 1. The beer-distribution game (adapted from Croson and Donohue, 2003)

- **POS data treatment** (treatment conducted by Croson and Donohue (2003): Includes data for 10 teams (40 players) placing orders for 48 weeks. All subjects know the distribution of final customer demand and observe final customer demand (POS data).
- **Inventory-data treatment** (treatment conducted by Croson and Donohue (2006):
  Includes data for 11 teams (44 players) placing orders for 48 weeks. All subjects know the distribution of final customer demand and observe other supply-chain players' weekly inventory data. Only retailers observe final customer demand (POS data).

Development of hypotheses

Subjects' orders: anchoring and adjustment heuristic

In the context of the beer game, Sterman (1989) proposed an anchoring and adjustment heuristic to explain the pattern of orders placed by subjects. His model suggests that subjects' orders account for (a) expected losses from the demand forecast; (b) an adjustment to maintain a desired level of inventory; and (c) an adjustment to maintain an adequate supply line of orders. He formalized the decision rule as:

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

where $\widehat{L}_t$ is the expected loss; S_t and SL_t the inventory and supply line at time t, respectively; S^* and SL^* the desired level of inventory and supply line; α_S and α_{SL} the fractional weekly inventory and supply-line adjustment; and constraints $\alpha_S, \alpha_{SL}, S^*, SL^* \geq 0$. Sterman modeled the expected loss with adaptive expectations: $\widehat{L}_t = \theta L_{t-1} + (1-\theta)\widehat{L}_{t-1}$, where $0 \leq \theta \leq 1$.

In Sterman's formulation (1), the *anchor* for the number of orders to be placed is given by the expected loss term ($\widehat{L}_t$); and the required *adjustment* maintains both inventory (S_t) and supply line (SL_t) at desired levels (S^*,SL^*). In this study, we decompose orders into their anchor and adjustment components and focus on the impact of information availability on each. Our decomposition approach is corroborated by Lawrence *et al.* (2006),

DOI: 10.1002/sdr

Schweitzer and Cachon (2000) and Kremer *et al.* (2011). In particular, Lawrence *et al.* (2006, p. 507) point to decomposition as a way to improve the accuracy of forecasts by “splitting the judgmental task into a series of smaller and cognitively less demanding tasks, and then combining the resulting judgments.” Schweitzer and Cachon (2000, p. 419) corroborate this suggesting that an “approach to improving inventory order decisions is to separate the forecasting task from the inventory decision task.” More recently, Kremer *et al.* (2011, p. 1838) hypothesize that “[d]ecomposing … inventory decisions may be a fruitful and important endeavor.”

Anchor decision

In a multiechelon serial supply chain with stationary demand and known distribution, such as the setting for Croson and Donohue’s (2003, 2006) beer game experiments, Chen (1999) determines that the optimal base-stock policy sets orders (O_t) as equal to those received in the previous period ($O_t = L_{t-1}$). Hence, a pass through policy, i.e. orders that are perfectly aligned with those received, is optimal in the beer game setting with stationary demand and known distribution. Prior research “suggest that people use the last data point in the series as a mental anchor and then adjust away from that anchor” (Harvey, 2007, p. 17). In the beer game, such results can be interpreted as the adoption of a naïve forecast ($\widehat{L}_t = L_{t-1}$), where the expected loss equals the previous period loss, i.e. the extreme case when $\theta = 1$ in an exponential smooth forecast ($\widehat{L}_t = \theta \cdot L_{t-1} + (1-\theta) \cdot \widehat{L}_{t-1}$). Accordingly, the expected coefficient for expected loss ($\widehat{L}_t$) should be 1; however, previous research finds lower coefficients, i.e. subjects do not fully account for expected loss in their ordering policy (Oliva and Gonçalves, 2007). Furthermore, research on decision support systems (DSS) suggest that additional information can help subjects make better decisions, outperforming unaided subjects (van Bruggen *et al.*, 1998). From this perspective, it is possible to conjecture that access to additional information (POS or inventory data) may allow subjects to generate better forecasts, improving their ability to explain the variance observed in oders placed.

In contrast, Oliva and Gonçalves (2007) document that most subjects (96 percent) in their beer game experiments use naive forecasting, that is, generate a forecast equal to the last demand signal observed. A naive forecast is also consistent with Harvey’s (2007, p. 17) last data-point mental anchor. From this alternative perspective, it is also possible to conjecture that even with additional information (POS or inventory data) subjects may still use a naïve forecast and not adjust orders based on it. In such cases, additional information (POS or inventory data) would not improve the forecast’s ability to explain the variance observed in oders placed.

DOI: 10.1002/sdr

Hypothesis 1A:

Availability of POS data will lead subjects to place orders that pay more attention to expected losses than orders placed without such information.

Hypothesis 1B:

Availability of inventory data will lead subjects to place orders that pay more attention to expected losses than orders placed without such information.

Adjustment decision

There are two components in the *adjustment* decision: inventory and supply line. In the sections below we address each one separately.

Inventory adjustment decision

When participants do not have access to additional relevant information (either POS or supply-chain partners' inventory data), their ordering decisions must be based solely on the available information cues (e.g. own inventory data, supply line, etc.). For instance, Lawrence and O'Connor (1995) find that, when testing an anchoring and adjustment model to real time-series data, subjects "anchor on the last history value of the series" and excessively adjust away from it (i.e. placing a strong emphasis on the adjustment component).

In addition, Kremer *et al*. finds that "subjects tend to overreact to observed forecast errors in relatively stable times series" (2011, p. 1838), when decision makers must face a succession of decisions (as in the beer game). In such cases, Kremer *et al*. suggest that subjects "may perform better when the relative salience of recent demand signals is mitigated, such as by reemphasizing the environment" (p. 1839). We postulate that access to additional information provides such a reemphasis of the environment, working to mitigate the "relative salience" of available inventory signals. Hence, we postulate that that access to additional information (either POS or inventory data) may allow participants to rely less on previously available information cues and incorporate the additional relevant information cues from the environment into their decisions. In particular, we expect participants to react less aggressively to their own inventory data, when they have access to POS or inventory data.

We empirically test these conjectures at an *aggregate level* (with data from all echelons) with the following formal hypotheses:

DOI: 10.1002/sdr

Hypothesis 2A:

Availability of POS data will lead subjects to place orders that react less aggressively to their own inventory than orders placed without such information.

Hypothesis 2B:

Availability of inventory data will lead subjects to place orders that react less aggressively to their own inventory than orders placed without such information.

Supply-line adjustment decision

One of Sterman (1989) key findings is that subjects' playing the beer game underweight the supply-line information. He postulates that supply-line underweighting is one of the behavioral causes that causes participants' ordering variation and Bullwhip Effect in the supply chain. Croson and Donohue (2003, 2006) confirm Sterman's (1989) finding that subjects' do not adjust their ordering decision with a supply-line adjustment. That result holds, even when subjects have access to POS information and supply-chain-partner inventory data. Since most subjects do not account for the supply-line adjustment, it is difficult to expect that the Lawrence and O'Connor (1995) prescription (e.g. a strong emphasis on a supply-line adjustment) would apply.

Since we use Croson and Donohue (2003, 2006) data in this study, we expect the supply-line underweighting bias to still be present when we estimate the coefficients of the model with a panel-data structure. However, since subjects have access to additional supply-chain information (e.g. POS or inventory data), we conjecture that subjects will have more concrete cues to infer their inventory position. POS data provide subjects with clear information on final customer orders and order variability, both of which reduce the cognitive load associated with their own ordering decisions. Inventory data across echelons provide subjects with supply-chain visibility, facilitating subjects inference on ability of suppliers to fulfill past (i.e. supply line) orders. While Croson and Donohue (2003, 2006) find that access to additional information (POS or inventory data) leads to orders with lower variation and reduced Bullwhip Effect, we conjecture that such effect is not only due to a smaller reaction to own inventory (4.3.1.) but also due to more attention on the supply line (4.3.2.). That is, we conjecture that access to additional information (POS or inventory data) allows subjects to place orders that pay more attention to the supply-line information:

DOI: 10.1002/sdr

Table 2. Key variables within two studies

Data source	Independent variable	Dependent variables
Croson and Donohue (2003)	Access to POS data	Expected loss coefficient Inventory coefficient Supply-line coefficient
Croson and Donohue (2006)	Access to inventory information	Expected loss coefficient Inventory coefficient Supply-line coefficient

Hypothesis 3A:

Availability of POS data will lead subjects to place orders that pay more attention to the supply line than orders placed without such information.

Hypothesis 3B:

Availability of inventory data will lead subjects to place orders that pay more attention to the supply line than orders placed without such information.

Table 2 summarizes the source of the data, independent and dependent variables tested in our study following our decomposition of the ordering decision into its anchor and adjustment components (Sterman, 1989) and the motivations for the possible impact of additional information (e.g. POS or inventory data) in each of them.

Relative impact of POS and inventory data

In the POS data treatment (Croson and Donohue, 2003), all 40 subjects know the distribution of final customer demand and observe final customer demand (POS data). In the inventory-data treatment (Croson and Donohue, 2006), all 44 subjects know the distribution of final customer demand. All retailers observe final customer demand (POS data). And, all subjects observe other supply-chain players' weekly inventory data. Comparing the POS and inventory-data treatments, we observe that in the latter subjects have access to dynamic inventory information which supposedly provides more information than POS data. In the inventory-data treatment all subjects still have information on the distribution of final (POS) customer demand. Hence, we conjecture that when subjects share inventory data in comparison with when they share POS data, they react less aggressively to their own inventory and pay more attention to supply-line information.

DOI: 10.1002/sdr

Given the ambiguous role of forecast in influencing order decisions, it is unclear how additional POS and inventory data will impact the forecast.

In the next three hypotheses, we conjecture that there are significant differences in the ordering behavior of participants when we increase the amount of information shared.

Hypothesis 4A:

Availability of inventory data across the supply chain will lead subjects to place orders that pay more attention to expected losses than orders placed with POS data.

Hypothesis 4B:

Availability of inventory data across the supply chain will lead subjects to place orders that react less aggressively to their own inventory than orders placed with POS data.

Hypothesis 4C:

Availability of inventory data across the supply chain will lead subjects to place orders that pay more attention to supply line than orders placed with POS data.

Econometric model and results

Econometric model

Our estimation model has its roots in Sterman's (1989) model using an anchoring and adjustment heuristic (1) to explain ordering behavior in the beer game. It departs from it in three significant ways. First, our model assumes a simple lag forecast ($\widehat{L}_t = L_{t-1}$) for the expected loss. That is, the demand forecast is given by the demand realization in the previous period. As mentioned earlier, in a multiechelon serial supply chain with known and stationary demand distribution, a pass-through policy, orders that are perfectly aligned with those received ($O_t = L_{t\text{-}1}$), is optimal. In the simplest case of the anchoring and adjustment heuristic, orders would be equal to the anchor, where the anchor would be given by the expected loss ($O_t = \widehat{L}_t = L_{t-1}$). In addition, the lag forecast is an intuitive model of expectation formation (Kleinmuntz, 1993) and has been used in previous empirical research with the beer game (Steckel *et al.*, 2004; Oliva and Gonçalves, 2007).

DOI: 10.1002/sdr

Second, consistent with Croson and Donohue (2003, 2006), our econometric model includes three other independent variables: inventory (S_t), supply line (SL_t), and received orders (R_t).[i] Such a model permits us to more readily compare our results with those of Croson and Donohue (2003, 2006).[ii]

Third, we structure the data as a panel—a cross-sectional time-series data set with individual players capturing the cross-sectional unit (i) and week of decision time index (t)—to estimate a decision rule that reflects the full range of observations available. Our methodological approach follows Oliva and Gonçalves (2007) and is in sharp contrast to a number of previous empirical studies (Sterman, 1989; Croson and Donohue, 2003, 2005, 2006) where researchers estimated decision rules for each individual subject. The panel-data structure increases the efficiency of the estimates and the representativeness of the resulting rule as it allows us to make estimations across individuals and echelons. Still, our model differs from Oliva and Gonçalves (2007). First, our model uses a fixed-effect, instead of a random-effects panel. The fixed-effect approach imposes a time-independent effect for each subject. This "fixed" effect focuses on the within-subject variation and manages to control for unmeasured covariates. Given the potential challenge associated with omitting other relevant variables, the fixed-effects method allows us to get unbiased estimates for the coefficients in our decision rule. Also, because in this study we are not particularly interested in subjects' reaction to backlog, we do not account for backlog explicitly as Oliva and Gonçalves (2007) did. Furthermore, we do not use a tobit model because "there is no command for a parametric conditional fixed-effects tobit model, as there does not exist a sufficient statistic allowing fixed effects to be conditioned out of the likelihood" (Stata Reference Manual Su-Z, p. 474, under "xttobit").

Introducing the variables for expected loss $\left(\widehat{L}_{it}\right)$, inventory ($S_{it}$), supply line ($SL_{it}$), and received orders ($R_{it}$) and the expansions for panel data and individual disturbances (α_i) (e.g. subject attitudes, cultural factors, abilities, and demographics) and ε_{it} (e.g. day of the week, time of day), yields the model:

$$O_{it} = \beta_0 + \beta_L L_{it-1} + \beta_S S_{it} + \beta_{SL} SL_{it} + \beta_R R_{it} + \alpha_i + \varepsilon_{it} \tag{2}$$

where S_{it} captures the inventory data for subject i at time t, and coefficients $\beta_0, \beta_L, \beta_S, \beta_{SL}, \beta_R$ capture the intercept and the fractional adjustment rates for expected loss, inventory, supply line, and received orders, respectively.

[i] Our panel-data structure, however, prevents us from including time (t) as an independent variable in our model.
[ii] We follow traditional research convention to include supply line and received orders as independent variables. However, we also present results of a model that does not include them, as both variables can be problematic. A model that includes supply line may suffer from endogeneity. While supply line influences orders (a high supply line should lead in lower orders, because the orders that have been placed will eventually be received), orders also directly affect the supply line (orders directly increase the supply line). Received orders may be collinear with supply line.

DOI: 10.1002/sdr

Aggregate level estimation results and analysis

Table 3 presents the estimation results for the fixed-effects panel model (2). The results show that our model is significant for all treatments. The adjusted R^2 for all three treatments is high, namely 0.59 in the baseline, 0.53 in the POS data treatment, and 0.47 in the inventory-data treatment.

The highly significant value of the F test allows us to reject the null hypothesis that the coefficients of the model are zero. In fact, most coefficients (e.g. constant (β_0), expected loss (β_L), inventory (β_S), and supply line (β_{SL})) are highly significant ($p < 0.001$), with the exception of the coefficient for received orders (β_R). In addition, most coefficients have the expected signs, with the exception of the supply line, which has a significant and positive coefficient. Croson and Donohue (2006) find a similar result with 20 of the 44 subjects in their base case showing positive supply-line coefficients. These results are also consistent with previous studies (Sterman, 1989; Croson and Donohue, 2006; Oliva and Gonçalves, 2007). The coefficient for expected loss (β_L) in the baseline treatment $\left(\beta_L^{Base} = 0.45\right)$ suggests that everything else held constant at a 10-percent increase in expected loss would lead to a 4.5-percent increase in orders placed. The coefficient for own inventory (β_S) in the baseline treatment $\left(\beta_S^{Base} = -0.12\right)$ suggests that everything else

Table 3. Aggregate level estimation results: baseline, POS, and inventory treatments $O_{it} = \beta_0 + \beta_L L_{it-1} + \beta_S S_{it} + \beta_{SL} SL_{it} + \beta_R R_{it} + \alpha_i + \varepsilon_{it}$

		Treatments		
	Regressor	Baseline	POS data	Inventory data
β_0	Constant	2.62***	2.06**	1.81**
		(0.34)	(0.68)	(0.65)
β_L	Expected Loss (L_{it-1})	0.45***	0.41***	0.42***
		(0.05)	(0.06)	(0.05)
β_S	Inventory (S_{it})	−0.12***	−0.07***	−0.08***
		(0.01)	(0.01)	(0.02)
β_{SL}	Supply Line (SL_t)	0.05***	0.11***	0.10***
		(0.01)	(0.03)	(0.02)
β_R	Received order (R_t)	−0.03	−0.01	0.08
		(0.03)	(0.03)	(0.06)
	F test	84.30***	150.57***	257.45***
	Adjusted R^2	0.59	0.53	0.47
	AIC	11,438	9412	11,122
	BIC	11,461	9435	11,144
	Observations	2112	1920	2112
	Number of players	44	40	44

Standard errors in parentheses.
*Significant with p < 0.05.
**p < 0.01.
***p < 0.001.

DOI: 10.1002/sdr

held constant at a 10-percent increase in own inventory would lead to a 1.2-percent decrease in orders placed. The coefficient for supply line (β_{SL}) in the baseline treatment $\left(\beta_S^{Base} = 0.05\right)$ suggests that everything else held constant at a 10-percent increase in the supply line would will lead to a 0.5-percent increase in orders placed.

Consider now the impact of additional information on the coefficients for expected loss (β_L). We observe that the coefficients for expected loss (β_L) change from $\left(\beta_L^B = 0.45\right)$ in the baseline treatment, to $\left(\beta_L^P = 0.41\right)$ in the POS treatment and to $\left(\beta_L^I = 0.42\right)$ in the inventory treatment. Availability of POS and inventory data reduce the strength of the coefficient (8.9- and 6.7-percent drop, respectively). However, considering the magnitude of the standard errors, the decrease and increase are not significantly different than the original result. Hence, the results suggest that availability of additional information (either POS or inventory) does not improve subjects' ability to place orders that are more aligned with those received. The results do not support Hypotheses 1A and 1B.

Consider now the impact of additional information on the coefficients for own inventory (β_S). We observe that the coefficients for own inventory change from $\beta_s^B = -0.12$ in the baseline treatment, to $\beta_S^P = -0.07$ in the POS treatment, and to $\beta_S^I = -0.08$ in the inventory treatment. Availability of POS and inventory data reduces the strength of the own inventory coefficient (a 33-percent drop) and the small magnitude of the standard errors suggests that the decrease is significant. Hence, the results suggest that with the availability of additional information (POS or inventory), subjects react less aggressively to their own inventory data, supporting Hypotheses 2A and 2B.

Consider next the impact of additional information on the coefficients for supply line (β_S). We observe that the coefficients for supply-line change from 0.05 in the baseline treatment, to 0.11 in the POS treatment, and to 0.10 in the inventory treatment. Availability of POS and inventory data increases the strength of the supply-line coefficient (a 50-percent increase), that is, it reduces supply-line underweighting. The small magnitude of the standard errors suggests that the increase is significant. Hence, the results suggest that with the availability of additional information (POS or inventory) subjects pay more attention to the supply line of previous orders placed, supporting Hypotheses 3A and 3B.

Aggregate level estimation results: POS and inventory treatment comparison

Comparing the values of the coefficients for expected loss (β_L) across information treatments, the results suggest a significant difference between the POS and the inventory-data treatments. In the POS treatment, the coefficient for expected loss $\left(\beta_L^P\right)$ is 0.41, and in the inventory-data treatment, the

DOI: 10.1002/sdr

coefficient $\left(\beta_L^I\right)$ is 0.42. Comparing the coefficients for expected loss (β_L) for the POS and the inventory-data treatments, our *t*-test ($t_{82} = 0.8$; $p = 0.78$) suggests that the coefficients are not significantly different. Thus, we cannot reject the null hypothesis that the individual-estimation coefficients for expected loss in the inventory-data treatment $\left(\beta_L^I\right)$ are lower than or equal to those in the POS treatment $\left(\beta_L^P\right)$ $\left(H_0 : \beta_L^I \le \beta_L^P\right)$. The results does not support Hypothesis 4A. Hence, we cannot claim that subjects place orders that are more aligned with those received when partners share inventory data across the supply chain instead of POS data.

Comparing the values of the coefficients for inventory across information treatments, our results again do not suggest a significant difference between the POS and the inventory-data treatments. In the POS treatment, the coefficient for inventory $\left(\beta_S^P\right)$ is 0.07, and in the inventory-data treatment the coefficient $\left(\beta_S^I\right)$ is 0.08. Comparing the coefficients for inventory (β_S), our *t*-test ($t_{82} = 3.4$; $p = 0.0015$) suggests that we can reject the null hypothesis that the individual-estimation coefficients for inventory in the inventory-data treatment $\left(\beta_S^{Inv}\right)$ are higher than or equal to those in the POS treatment $\left(\beta_L^{POS}\right)$ $\left(H_0 : \beta_S^{Inv} \ge \beta_S^{POS}\right)$. The result supports Hypothesis 4B, that availability of inventory data across the supply chain will lead subjects to place orders that react less aggressively to their own inventory than orders placed with POS data. Comparing the values of the coefficients for supply line across information treatments, our results again do not suggest a significant difference between the POS and the inventory-data treatments. In the POS treatment, the coefficient for supply line β_{SL}^P is 0.11; and in the inventory-data treatment, the coefficient β_{SL}^I is 0.10. Comparing the coefficients for supply line β_{SL}, our *t*-test ($t_{82} = 1{,}38$; $p = 0.09$) suggests that we cannot reject the null hypothesis that the individual-estimation coefficients for supply line in the inventory-data treatment β_{SL}^I are higher than or equal to those in the POS treatment β_{SL}^P $\left(H_0 : \beta_{SL}^I \ge \beta_{SL}^P\right)$. The results do not support Hypothesis 4C, that availability of inventory data across the supply chain will lead subjects to place orders that pay more attention to supply line than orders placed with POS data.

We formally tested these results with dummy variables (see Appendix A.3, Appendix S1), and in addition assessed the robustness of our model by running a number of tests to check the validity of our assumptions (see Appendix A.4, Appendix S1 in the online supporting information).

Upstream-level estimation results and analysis

So far, we have analyzed the aggregate data (for all positions) in the beer game. However, Croson and Donohue (2003, 2006) distinguish between the behavior of upstream and downstream players in the game. A number of

DOI: 10.1002/sdr

scholars (e.g. Bourland *et al*., 1996; Gavirneni *et al*., 1999; Cachon and Fisher, 2000; Croson and Donohue, 2005) argue that upstream members benefit more than the downstream members from information sharing in supply chains. Chen (1999) maintains that access to information from other supply chain members helps *manufacturers* and *distributors* better handle their decisions. Croson and Donohue (2003, 2005, 2006) find that inventory data and POS data reduce order oscillation where *upstream* echelons seem to make better use of that additional information. In particular, Croson and Donohue conclude that "inventory information … improves performance by allowing manufacturers and distributors to anticipate and interpret orders placed by their downstream customers" (2006, p. 333). The next set of analyses (Table 4), consider our model results when we divide players into upstream (manufacturer and distributor) and downstream (wholesaler and retailer) groups. Below, we present the analyses for upstream player; Appendix 4, Appendix S1 in the online supporting information presents the analysis for downstream ones.

Consider the coefficients for expected loss β_L^U for upstream players; they change from $\beta_L^{UB} = 0.46$ in the baseline treatment to $\beta_L^{UP} = 0.46$ in the POS treatment and to $\beta_L^{UI} = 0.42$ in the Inventory treatment. Availability of POS data does not change the strength of the coefficient, but the availabilty of inventory data decreases the strength of the coefficient (9-percent drop). Considering the coefficients for own inventory β_S^U, we observe that they change from $\beta_S^{UB} = -0.12$ in

Table 4. Upstream- and downstream-level estimation results for each treatment

		Treatments					
		Baseline		POS data		Inventory data	
	Regressor	Downstream	Upstream	Downstream	Upstream	Downstream	Upstream
β_0	Constant	3.27***	2.64***	4.31***	1.34***	3.33***	1.32
β_L	Expected Loss (L_{t-1})	0.32***	0.46***	0.28**	0.46***	0.31***	0.42***
β_S	Inventory (S_t)	−0.11***	−0.12***	−0.15***	−0.05***	−0.12***	−0.05*
β_{SL}	Supply Line (SL_t)	0.03*	0.09*	0.03*	0.15**	0.04*	0.15***
β_R	Received order (R_t)	−0.09	−0.06	−0.04	−0.02	0.04	0.07
	F test	100.07***	141.26***	118.68***	139.93***	72.84***	221.80***
	Adjusted R 2	0.48	0.64	0.49	0.58	0.38	0.53
	AIC	5261	5990	4305	4921	5225	5763
	BIC	5280	6010	4325	4940	5245	5783
	Observations	1056	1056	960	960	1056	1056
	Number of players	22	22	20	20	22	22

*Significant with $p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

DOI: 10.1002/sdr

the baseline treatment, to $\beta_S^{UP} = -0.05$ in the POS treatment, and to $\beta_S^{UI} = -0.05$ in the inventory treatment. Availability of POS and inventory data reduces the strength of the own inventory coefficient (a 58-percent drop).

Considering the coefficients for supply-line β_{SL}^{U}, we observe that they change from $\beta_{SL}^{UB} = 0.09$ in the baseline treatment, to $\beta_{SL}^{UP} = 0.15$ in the POS treatment, and to $\beta_{SL}^{UI} = 0.15$ in the inventory treatment. Availability of POS and inventory data increases the strength of the supply-line coefficient (a 67-percent gain). Additional information (POS or inventory data) causes upstream players to increase their reaction to expected losses and supply line and to decrease their reaction to their own inventory.

To formally test Hypotheses 1A/1B, 2A/2B, and 3A/3B for upstream players, we ran the model with dummy variables for the POS treatment (*P*) and inventory treatment (*I*) interacted with the coefficients of interest (Table 5).

Table 5. Echelon-level estimation results

		Treatments					
		Baseline		POS data		Inventory	
	Regressor	Downstream	Upstream	Downstream	Upstream	Downstream	Upstream
β_0	Constant	3.27****	2.64****	3.77****	2.02****	3.30****	1.98****
β_L	Expected Loss (L_{t-1})	0.32****	0.46****	0.32****	0.46****	0.32****	0.46****
β_S	Inventory (S_t)	−0.11****	−0.12****	−0.11****	−0.12****	−0.11****	−0.12****
β_{SL}	Supply Line (SL_t)	0.03**	0.09***	0.03**	0.09****	0.03***	0.09**
β_R	Received order (R_t)	−0.09**	−0.06	−0.09**	−0.06	−0.09**	−0.06
β_0^P	*P* (POS dummy)			(omitted)	(omitted)		
β_L^P	PL_{t-1}			−0.04	0.00		
β_S^P	PS_t			−0.04*	0.07***		
β_{SL}^P	PSL_t			0.00	0.06		
β_R^P	PR_t			0.05	0.04		
β_0^I	*I* (Inventory dummy)					(Omitted)	(Omitted)
β_L^I	IL_{t-1}					−0.01	−0.04
β_S^I	IS_t					0.00	0.07**
β_{SL}^I	ISL_t					0.01	0.06*
β_R^I	IR_t					0.13**	0.12
	F test	100.07****	141.26****	112.15****	144.13****	88.51****	185.85****
	Adjusted R 2	0.48	0.64	0.48	0.62	0.44	0.60
	AIC	5261	5990	9628	10,985	10,486	12,064
	BIC	5280	6010	9672	11,030	10,532	12,086
	Observations	1056	1056	2016	2016	2112	2112
	Number of players	22	22	42	42	44	44

*Significant with p < 0.10.
**p < 0.05.
***p < 0.01.
****p < 0.001.

DOI: 10.1002/sdr

Considering the results for the interaction coefficients for expected loss (β_L^{UP} and β_L^{UI}), the *p-value* results ($p_L^{UP} = 0.96; p_L^{UI} = 0.71$) confirm that the upstream-level data does not support Hypotheses 1A and 1B. Considering the results for the interaction coefficients for own inventory (β_S^{UP} and β_S^{UI}), the *p-value* results ($p_S^{UP} = 0.001; p_S^{UI} = 0.013$) confirm that the upstream-level data supports Hypotheses 2A and 2B. We find that additional information *directionally* helps upstream subjects place orders that are less reactive to their own inventory data (an average 58-percent decrease). Considering the results for the interaction coefficients for supply line $\left(\beta_{SL}^{UP}\, and\, \beta_{SL}^{UI}\right)$, the *p-value* results $\left(p_{SL}^{UP} = 0.242; p_{SL}^{UI} = 0.085\right)$ confirm that the upstream-level data does not support Hypothesis 3A, but it supports Hypothesis 3B ($p < 0.1$).

Discussion

Our research compiles and analyzes data from two beer game experiments (Croson and Donohue, 2003, 2006) to gain insight on the specific mechanisms that reduce order amplification. Structuring the data as a panel and using a fixed-effect model enabled us to efficiently obtain an unbiased and representative estimate of how additional information such as POS data or supply chain partners' inventory data affects orders placed by subjects.

Aggregate-level results

Table 6 summarizes our aggregate-level results for the four proposed hypotheses. First, while it is optimal for subjects in the beer game to adopt a "pass-through" ordering policy (i.e. order the same amount as orders received),

Table 6. Summary of results from examined hypotheses at an aggregate level

Hypothesis	Effect of	On	Results
H1A	Availability of POS data	Expected loss	Not Supported
H1B	Availability of inventory data	Expected loss	Not Supported
H2A	Availability of POS data	Inventory	Supported
H2B	Availability of inventory data	Inventory	Supported
H3A	Availability of POS data	Supply line	Supported
H3B	Availability of inventory data	Supply line	Supported
H4A	Availability of inventory data compared with availability of POS data	Expected loss	Not Supported
H4B	Availability of inventory data compared with availability of POS data	Inventory	Supported
H4C	Availability of inventory data compared with availability of POS data	Supply line	Not Supported

DOI: 10.1002/sdr

they account for only a fraction of expected losses ($\widehat{L}_t$) in their orders. Furthermore, our results suggest that access to additional information (POS or inventory data) does not significantly affect subjects' use of forecasts in their own order policy (no support for *H1*). Second, our research indicates that additional information (POS or inventory data) causes subjects to be more conservative in eliminating inventory gaps, that is, they react less aggressively to their own inventory data (support for *H2*). The estimated coefficient for own inventory in the baseline condition was -0.12 meaning that, all else equal, a 1-percent change in the value of own inventory leads to a 12-percent decrease in the order amount placed. In the POS and inventory-information treatments, the estimated coefficient for own inventory were -0.07 and -0.08 meaning that, all else equal, a 1-percent change in the value of own inventory leads to only a 7-percent decrease and an 8-percent decrease in orders placed, respectively. That is, a statistically significant lower reaction of own inventory on orders placed.

Third, additional information causes subjects to pay more attention to the supply line (support for *H3*). The estimated coefficient for supply line in the baseline condition was $+0.05$ meaning that, all else equal, a 1-percent change in the value of the supply line leads to a 5-percent increase in the order amount placed. In the POS and inventory-information treatments, the estimated coefficient for own inventory were $+0.11$ and $+0.10$ meaning that, all else equal, a 1-percent change in the value of the supply line would lead to 10- and 11-percent increase, respectively, in orders placed. That is, subjects have a statistically significant stronger reaction to the supply-line information in the treatment conditions than on the baseline condition. So, subjects pay more attention to the supply line which translates into a higher magnitude of the supply-line impact.

Fourth, analysis on the comparative impact of POS and inventory information suggest that both have a similar noneffect on expected losses (no support for *H4A*), both have a similar effect on the supply line (no support for *H4C*), and inventory information has a stronger effect on own inventory (support for *H4B*).

Croson and Donohue (2003, 2006) find that additional information (POS data and supply-chain partners' inventory) decreases the Bullwhip Effect in the supply chain; our analysis focuses on possible mechanisms affecting subjects orders that could explain this decrease. Our analyses indicate that availability to POS and inventory information affects subjects' orders in two ways: (1) a reduced reaction to own inventory information, that is, subjects react less aggressively to their own inventory data (*H2*); and (2) reduced underweighting of the supply line, that is, subjects pay more attention to the supply line (*H3*).

Upstream-level results

First, our estimation results at the echelon level show that POS and inventory information do not significantly impact either downstream or upstream

DOI: 10.1002/sdr

subjects to place orders that are more aligned with those received than those placed (no support for *H1* at the echelon level).[iii] Second, our results show that POS and inventory data cause upstream players to be less reactive to their own inventory information (support for *H2* at the echelon level for upstream players). Third, additional POS and inventory information cause upstream players to pay more attention to the supply line of previous orders placed (support for *H3A* at the echelon level for upstream players). In summary, all upstream players seem to use the additional information to react less to their own inventory and to pay more attention to the supply line. Our findings are well aligned with previous hypotheses for upstream players, such as the ability to "better anticipate [downstream] customers' needs" (Croson and Donohue, 2003, p. 10) and "improve a decision maker's ability to anticipate supply shortages… [and] combat the supply-line underweighting tendency" (Croson and Donohue, 2006, p. 330).

Managerial implications

These results have implications for supply-chain partners. As Croson and Donohue (2006, p. 334) speculate, inventory sharing brings more value when they track and share "inventory position of the retailer to the manufacturer." Such information helps upstream players (i.e. manufactures and distributors) react less aggressively to their own inventory data and pay more attention to the supply line. The observed change in behavior is more pronounced in upstream echelons, which suggests that manufactures and distributors gain more through sharing POS or inventory data along the supply chain. A major benefit of sharing POS or supply-chain inventory information is a reduced inventory cost. Since, upstream members reap the bulk of the benefits, it is natural to expect that they bear most of the expense of setting up such information-sharing systems. It is also to their advantage to design mechanisms (or policies) that motivate downstream members to share their POS and inventory information.

Comparative analysis to other studies

Table 7 compares coefficient estimates in our study with several others (e.g. Sterman, 1989; Croson and Donohue, 2002, 2003, 2005, 2006; Oliva and Gonçalves, 2007, Croson *et al*., 2014; Sterman and Dogan, 2015; Oliva *et al*., 2021). Most of the studies have similar results. Our findings conform with that of others with respect to the coefficients for expected loss and own inventory. It differs from some results with respect to estimated coefficients for the supply line of past orders. Below, we discuss the similarities and important differences.

[iii] Appendix A.4.1. in the online supporting information presents the results for downstream players.

DOI: 10.1002/sdr

Table 7. Comparative results from key behavioral studies using the beer game

Study	Expected Loss (L_{it-1})	Inventory (S_{it})	Supply line (SL_t)	Backlog (S_tB_t)	Delivery time (κ, ω)	Notes
Sterman (1989)[a]	0.36	−0.26	−0.09[a]			
Croson and Donohue (2002)[a]	N/A	−0.22	−0.03[a]			
Croson and Donohue (2003)[a]	0.33	−0.24	−0.03[a]			Baseline
	0.36	β_I	β_{SL}			POS information Treatment ($\beta_{SL} > \beta_I$)
Croson and Donohue (2005)[a]	N/A	N/A	N/A			N/A: Coefficient estimates not available
Croson and Donohue (2006)[a]	0.33	−0.24	−0.03[a]			Baseline
	0.33	−0.19	−0.03[a]			Inventory information treatment
Oliva and Gonçalves (2007)[b]	1.00[e]	−0.11	−0.01[f]			Aggregate Model I
	1.00[e]	−0.21	+0.00[f]	+0.21		Aggregate Model VI
Croson *et al*. (2014)[a]	0.09	−0.47	−0.04[c]			Baseline (T1)
	0.08	−0.28	−0.03[c]			Common knowledge (T2)
	0.00	−0.24	−0.04[c]			Eliminating coordination risk (T3)
	0.08	−0.26	−0.06[c]			Coordination stock (T4)
Sterman and Dogan (2015)[a]	0.32[d]	−0.32	−0.12[c]			Model 0
	0.38[d]	−0.33	−0.07[c]		22.0; 2.5	Model 1 Best
	0.36[d]	−0.23	−0.02[c]		21.5; 3.9	Model 2 Best
	0.37[d]	−0.28	−0.05[c]		21.7; 3.2	Model 1 or 2 Best
Oliva *et al*. (2021)[b]	1.00[e]	−0.11	−0.01[f]			Aggregate Model 1
	1.04	−0.21	+0.01[f]	+0.22		Aggregate Model 4
This Study	0.45	−0.12	+0.05			Baseline
	0.41	−0.07	+0.11			POS information Treatment
	0.42	−0.08	+0.10			Inventory information treatment

aValues for coefficients of Expected Loss (L (it-1)), Inventory (Sit), and Supply Line (SLt) reflect reported median values.
bValues for coefficients reflect aggregate values for all players estimated with panel data.
cValues indicate the SL coefficient given by the product of reported α and β.
dCoefficients for Expected Loss (L(it-1)) in Sterman and Dogan (2015) reflect reported median values of θ.
eAssumes lag forecast has coefficient equal to 1.
fCoefficient not significant at 10% level.

DOI: 10.1002/sdr

First, the estimated coefficient for expected losses in the baseline condition is 0.45, a value that has the same sign and has a similar magnitude to the 0.36 mean estimate in Sterman (1989) and the 0.33 in Croson and Donohue (2003, 2006). The estimated coefficient for own inventory in the baseline condition is −0.12, a value that has the expected sign and a magnitude lower than some studies (e.g. Sterman, 1989; Croson and Donohue, 2002, 2003, 2006; Croson *et al.*, 2014; Sterman and Dogan, 2015) and equal to others (Oliva and Gonçalves, 2007, and Oliva *et al.*, 2021). The estimated coefficient for supply line in the baseline condition is +0.05, a value that has an unexpected sign but of similar magnitude in comparison with most other studies. While a positive and significant effect of supply line is counterintuitive, it has been reported in previous published studies particularly for estimations of individual nonfactory players (e.g. Oliva and Gonçalves, 2007; Oliva *et al.*, 2021). In particular, Oliva and Gonçalves (2007) find positive ($\beta_{SL} = +0.02$) and statistically significantly (1 percent) supply-line coefficients estimating aggregate nonfactory positions. Oliva *et al.* (2021) find positive but not statistically significantly at 10 percent estimating aggregate supply-line coefficients for all positions. Importantly, Sterman and Dogan (2015) capture two feedback processes going through the supply line of orders with the supplier (the channel-order backlog): one balancing and one reinforcing (Figure 2)[iv]. The Supply-Line-Control balancing loop (*B2*) "close[s] any gap between the desired and actual supply line of goods on order with the supplier." Econometrically, *B2* would capture a negative coefficient of the supply line on orders. In the Phantom Orders reinforcing loop (*R1*), the supply line of orders causes delivery times to rise as "allocations fall and delivery reliability drops" leading to higher orders. Econometrically, *R1* would capture a positive coefficient of the supply line on orders (going through delivery times). Since most beer game studies do not capture the coefficient for delivery times on orders, econometrically, they estimate the net effect of the direct pathway of supply line to orders (in loop *B2*) and the indirect pathway of supply line to orders through delivery times (in loop *R1*). Theoretically, the sign of the supply-line coefficient can either be positive or negative depending on the relative strengths of those pathways. Oliva and Gonçalves (2007) and this study find a positive supply-line coefficient. Sterman and Dogan (2015) find large and positive expected delivery times ($\lambda^e = min(\lambda^M, \kappa + \omega\lambda^p)$) characterizing a strong gain for loop *R1*, a result that is well aligned with a positive and significant overall impact of the supply line on orders.

Prior studies (Sterman, 1989; Diehl and Sterman, 1995; Croson and Donohue, 2003, 2006; Croson *et al.*, 2014) assume that the Desired Supply Line (SL^*) and the Desired Inventory (S^*) are constant, effectively cutting

[iv] They describe a third feedback process, a reinforcing loop capturing the impact of hoarding (R2), whereby the supplier lead time influences the desired inventory level.

DOI: 10.1002/sdr

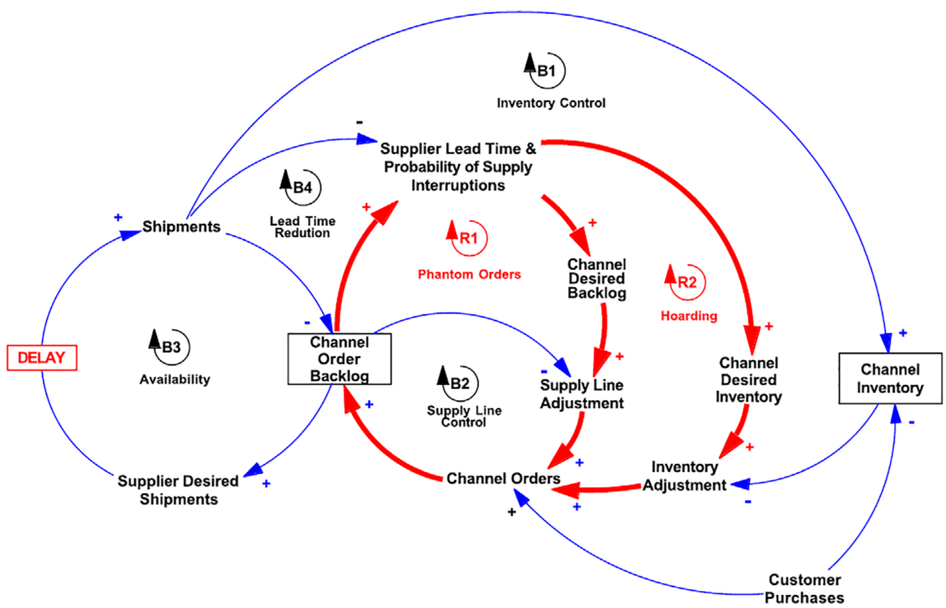

Fig. 2. Feedback structure of hoarding and phantom ordering (Sterman and Dogan, 2015) [Color figure can be viewed at wileyonlinelibrary.com]

reinforcing feedback loops R1 and R2, and rulling out phantom ordering and hoarding, respectively. In contrast, Sterman and Dogan (2015) assume those loops to be active and the Desired Supply Line (SL^*) and the Desired Invetory (S^*) to be dynamic and described by the following equations:

$$S^* = \gamma D_t^e, \tag{3}$$

$$SL^* = \lambda^e R_t^* = \lambda^e D_t^e, \tag{4}$$

where, the Desired Inventory (S^*) is given by the product of some desired-inventory coverage (γ), proportional (δ_1) to the actual delivery delay (λ), and the expected incoming orders (D_t^e). The Desired Supply Line (SL^*) is given by the product of the desired delivery rate (R^*), assumed equal to the expected incoming orders (D_t^e) in model #1, and the expected delivery delay (λ^e), proportional δ_2 to the actual delivery delay (λ). Substituting terms on Sterman's (1989) anchoring and adjustment heuristic for subjects' orders, we obtain:

$$O_t = D_t^e + \alpha_S(S^* - S_t) + \alpha_{SL}(SL^* - SL_t), \tag{5}$$

$$O_t = D_t^e + \alpha_S\left(\gamma D_t^e - S_t\right) + \alpha_{SL}\left(\lambda^e D_t^e - SL_t\right), \tag{6}$$

$$O_t = D_t^e + \beta\lambda D_t^e - \alpha_S S_t - \alpha_{SL} SL_t, \text{where } \beta = \alpha_S\delta_1 - \alpha_{SL}\delta_2. \tag{7}$$

Equation (7) provides a simple formula capturing subjects' ordering policies that incorporates endogenous estimates for the Desired Supply Line (SL^*) and

DOI: 10.1002/sdr

the Desired Invetory (S^*), while also disentangling the effects of the supply line. Equation (7) could be potentially estimated by keeping track of the actual delivery delay (λ) experienced by subjects and their expected incoming orders (D_t^e) and estimating the coefficient for the multiplicative term (λD_t^e).

These results from previous research suggest that our finding cannot be dismissed as due to a misspecification of our model. Instead, they suggest that our finding of a positive supply-line coefficient is present in previous research, and it should be taken into consideration.

Limitations and future research

This study has several limitations. First, the analysis is conducted in aggregate and at upstream/downstream levels. Were we to collect more observations per role, expected incoming orders, actual delivery delays, we may gain further insight on the behavior of roles throughout the supply chain. For example, there is an opportunity for further research clarifying how downstream players could use POS and inventory information to improve their ordering decisions. Second, the examined datasets were collected through previously conducted experiments; thus we did not have the opportunity to deepen our findings through post experiments debriefing interviews with subjects.

Researchers in behavioral-operations management recognize the value that decomposing a decision rule (e.g. separating forecasting from the inventory decision or separating the anchor and adjustment decisions) may have in improving inventory-ordering performance. While Lawrence *et al.* recognize that decomposition may be helpful, they lament that there has been surprisingly "little research over the last 25 years into the value of decomposition and the conditions under which it is likely to improve accuracy" (2006, p. 508). Our research attempts to decompose the ordering decision in the beer game into several precise experiments that focus on specific mechanisms that can help subjects place orders.

Given the impact of Bullwhip Effect on supply chains, scholars have called for more research to investigate the behavior of actors within a supply chain (e.g. Bolton and Katok, 2008; Narayanan and Moritz, 2015). Sterman and Dogan (2015) assert that behavioral-operations research can bridge the gap between traditional operations research and management with other behavioral sciences such as psychology, neuroscience, and organizational science to provide insight into supply-chain dynamics and deliver impactful suggestions to managers.

Acknowledgements

We are grateful to Rachel Croson and Karen Donohue for providing the data analyzed here and for their helpful comments. We also thank the referees

DOI: 10.1002/sdr

and associated editor who provided useful suggestions to improve the manuscript. All errors remain ours.

Biographies

Paulo Gonçalves is Professor of Management and Director of the Humanitarian Operations Group at the Università della Svizzera italiana (USI), Switzerland. He is also Research Fellow at the University of Cambridge Judge Business School (CJBS). His research combines system dynamics simulation, behavioral experiments, optimization, and econometrics, to understand how managers make strategic, tactical and operational decisions in humanitarian settings. Paulo holds a Ph.D. in Management Science and System Dynamics from the MIT Sloan School of Management and a M.Sc. from the Massachusetts Institute of Technology (MIT).

Mohammad Moshtari is an Academy of Finland Research Fellow, and Associate Professor of Supply Chain Management at Tampere University. His research centres on supply chain management in private and public sectors including NGOs and Higher Education Institutes, humanitarian operations and crisis management, and operations in emerging markets.

References

Bolton GE, Katok E. 2008. Learning by doing in the newsvendor problem: a laboratory investigation of the role of experience and feedback. *Manufacturing & Service Operations Management* **10**(3): 519–538.

Bourland KE, Powell SG, Pyke DF. 1996. Exploiting timely demand information to reduce inventories. *European journal of operational research*, **92**(2), 239–253.

Cachon GP, Randall T, Schmidt GM. 2007. In search of the bullwhip effect. *Manufacturing and Service Operations Management* **9**(4): 457–479.

Chen F. 1999. Decentralized Supply Chains Subject to Information Delays. *Management Science* **45**(8): 1076–1090.

Croson R, Donohue K. 2002. Experimental economics and supply-chain management. *Interfaces* **32**(5): 74–82.

Croson R, Donohue K. 2003. Impact of POS data sharing on supply chain management: an experimental study. *Production and Operations Management* **12**(1): 1–11.

Croson R, Donohue K. 2005. Upstream versus downstream information and its impact on the bullwhip effect. *System Dynamics Review* **21**(3): 249–260.

Croson R, Donohue K. 2006. Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science* **52**(3): 323–336.

DOI: 10.1002/sdr

Croson R, Donohue K, Katok E, Sterman J. 2014. Order stability in supply chains: coordination risk and the role of coordination stock. *Production and Operations Management* **23**(2): 176–196.

Cachon GP, Fisher M. 2000. Supply chain inventory management and the value of shared information. *Management science*, **46**(8), 1032–1048.

Chen F, Drezner Z, Ryan JK, Simchi-Levi D. (2000) Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management science*, **46**(3), 436–443.

Diehl E, Sterman JD. 1995. Effects of feedback complexity on dynamic decision making. *Organizational behavior and human decision processes*, **62**(2), 198–215.

Duc TTH, Luong HT, Kim YD. 2008. A measure of bullwhip effect in supply chains with a mixed autoregressive-moving average demand process. *European Journal of Operational Research*, **187**(1), 243–256.

Gavirneni S, Kapuscinski R, Tayur S. 1999. Value of information in capacitated supply chains. *Management science*, **45**(1), 16–24.

Harvey N. 2007. Use of heuristics: insights from forecasting research. *Thinking Reasoning* **13**(1): 5–24.

Khan MH, Ahmed S, Hussain D. 2019. Analysis of bullwhip effect: a behavioral approach. *Supply Chain Forum: An International Journal*, **20**(4), 310–331. https://doi.org/10.1080/16258312.2019.1661756.

Kleinmuntz DN. 1993. Information processing and misperceptions of the implications of feedback in dynamic decision making. *System Dynamics Review* **9**(3): 223–237.

Kremer M, Moritz B, Siemsen E. 2011. Demand forecasting behavior: system neglect and change detection. *Management Science* **57**(10): 1827–1843.

Lawrence M, O'Connor M. 1995. The anchor and adjustment heuristic in time-series forecasting. *Journal of Forecasting* **14**(5): 443–451.

Lawrence M, Goodwin P, O'Connor M, Önkal D. 2006. Judgmental forecasting: a review of progress over the last 25 years. *International Journal of Forecasting* **22**(3): 493–518.

Lee H, Padmanabhan V, Whang S. 1997. Information distortion in a supply chain: the bullwhip effect. *Management Science* **43**(4): 546–558.

Luong HT 2007. Measure of bullwhip effect in supply chains with autoregressive demand process. *European Journal of Operational Research*, **180**(3), 1086–1097.

Macdonald JR, Frommer ID, Karaesmen IZ. 2013. Decision making in the beer game and supply chain performance. *Operations Management Research* **6**(3–4): 119–126.

Narayanan A, Moritz BB. 2015. Decision making and cognition in multi-echelon supply chains: an experimental study. *Production and Operations Management* **24**(8): 216–1234.

Oliva, Abdulla, Goncalves, 2021. Do Managers Overreact When in Backlog? Evidence of Scope Neglect from a Supply Chain Experiment Texas A&M University. Working paper.

Oliva R, Gonçalves P. 2007. Behavioral causes of the bullwhip effect:"satisficing" policies with limited information cues. Mays School of Business Working Paper.

Schweitzer ME, Cachon G. 2000. Decision bias in the newsvendor problem with a known demand distribution: experimental evidence. *Management Science* **46**(3): 404–420.

DOI: 10.1002/sdr

Steckel JH, Gupta S, Banerji A. 2004. Supply chain decision making: will shorter cycle times and shared point-of-sale information necessarily help? *Management Science* **50**(4): 458–464.

Sterman J. 1989. Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes* **43**(3): 301–335.

Sterman JD, Dogan G. 2015. "I'm not hoarding, I'm just stocking up before the hoarders get here.": behavioral causes of phantom ordering in supply chains. *Journal of Operations Management* **39**: 6–22.

Sethuraman K, Tirupati D. 2005. Evidence of bullwhip effect in healthcare sector: causes, consequences and cures. *International Journal of Services and Operations Management* **1**(4): 372–394.

van Bruggen G, Smidts A, Wierenga B. 1998. Improving decision making by means of a marketing decision support system. *Management Science* **44**(5): 645–658.

Villa S, Gonçalves P, Arango S. 2015. Exploring retailers' ordering decisions under delays. *System Dynamics Review* **31**(1–2): 1–27.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's website.

Appendix S1. Supporting Information.

DOI: 10.1002/sdr