“I’m not hoarding, I’m just stocking up before the hoarders get here.”
Behavioral causes of phantom ordering in supply chains

John D. Sterman\textsuperscript{a,}\footnote{Corresponding author.\newline E-mail address: jsterman@mit.edu (J.D. Sterman).}, Gokhan Dogan\textsuperscript{b}
\textsuperscript{a} MIT Sloan School of Management, Cambridge, MA, USA\textsuperscript{b} TAV, Istanbul, Turkey

\begin{abstract}
When suppliers are unable to fill orders, delivery delays increase and customers receive less than they desire. Customers often respond by seeking larger safety stocks (hoarding) and by ordering more than they need to meet demand (phantom ordering). Such actions cause still longer delivery times, creating positive feedbacks that intensify scarcity and destabilize supply chains. Hoarding and phantom ordering can be rational when customers compete for limited supply in the presence of uncertainty or capacity constraints. But they may also be behavioral and emotional responses to scarcity. To address this question we extend Croson et al.’s (2014) experimental study with the Beer Distribution Game. Hoarding and phantom ordering are never rational in the experiment because there is no horizontal competition, randomness, or capacity constraint; further, customer demand is constant and participants have common knowledge of that fact. Nevertheless 22\% of participants place orders more than 25 times greater than the known, constant demand. We generalize the ordering heuristic used in prior research to include the possibility of endogenous hoarding and phantom ordering. Estimation results strongly support the hypothesis, with hoarding and phantom ordering particularly strong for the outliers who placed extremely large orders. We discuss psychiatric and neuroanatomical evidence showing that environmental stressors can trigger the impulse to hoard, overwhelming rational decision-making. We speculate that stressors such as large orders, backlogs or late deliveries trigger hoarding and phantom ordering for some participants even though these behaviors are irrational. We discuss implications for supply chain design and behavioral operations research.
\end{abstract}

\section{Introduction}
During World War II, as the Allies faced shortages of food and basic goods, hoarding became a serious threat. A cartoon at the time showed a stern store manager confronting a shopper attempting to buy dozens of cans of food despite rationing. Caught red-handed, the shopper says, “I’m not hoarding. I’m just stocking up before the hoarders get here.” Is such behavior a rational, if anti-social, response to scarcity, or an emotional reaction driven by fear and panic?

Hoarding, defined here as attempts to accumulate large private stocks of goods when people perceive threats to supply, is closely related to phantom ordering in which people react to uncertain supply by ordering more than they actually desire, or ordering from multiple suppliers, then planning to cancel their excess orders once they get what they desire. Not limited to wartime, hoarding and phantom ordering remain persistent, destabilizing and costly phenomena in supply chains. For example, during the great technology boom of the late 20th century, firms such as Cisco Systems, Lucent, Nortel, and JDS Uniphase experienced huge surges in incoming orders. Deliveries could not keep pace. Customers were placed on allocation, receiving only a fraction of what they ordered. Desperate for product, many customers ordered still more, often placing orders through multiple channels—in some cases, three or more times the number of units they actually desired (Goetz, 2005). These phantom orders further inflated backlogs, causing still longer delivery times and smaller allocations, a positive feedback that intensified scarcity. After a lag, production increased and allocations were eased. Suddenly able to get all the product they wanted, customers cancelled their phantom orders, leaving suppliers with huge excess stocks, excess capacity and deep losses. Cisco was forced to write off $2.2 billion in excess inventory. Others fared far worse. During the boom, the order backlog of equipment maker JDS Uniphase exploded, rising 3000\% from mid 1998 to mid 2000 (Fig. 1). Sales quadrupled between the end of 1999 and beginning of 2001. Uniphase expanded capacity and employment dramatically.

Please cite this article in press as: Sterman, J.D., Dogan, G., “I’m not hoarding, I’m just stocking up before the hoarders get here.”. Behavioral causes of phantom ordering in supply chains. J. Operations Manage. (2015). http://dx.doi.org/10.1016/j.jom.2015.07.002

http://dx.doi.org/10.1016/j.jom.2015.07.002
0272-6963/© 2015 Elsevier B.V. All rights reserved.
As output grew and product became available, new orders dried up and customers cancelled orders. Backlog collapsed, and sales fell 83% by the end of 2002. Uniphase cut employment by more than 23,000 (81%) and saw its stock price fall 99%.

Prior research offers two categories of explanation for hoarding and phantom ordering, and supply chain instability generally: operational and behavioral. Operational theories focus on the physical and institutional structure of supply chains, while assuming that decision makers are rational agents who make optimal decisions given their local information and incentives. Physical structure includes the network linking customers and suppliers and the placement of inventories and buffers within it, along with capacity constraints and time delays in production, order fulfillment, transportation, and so on. Institutional structure includes the degree of horizontal and vertical coordination and competition among firms, the availability of information, and the incentives faced by each decision maker. Behavioral explanations also capture the physical and institutional structure of supply chains, but view people as boundedly rational actors with imperfect mental models of the environment who use heuristics to make decisions (Morecroft, 1985; Sterman, 2000; Boudreau et al., 2003; Gino and Pisano, 2008; Bendoly et al., 2010; Croson et al., 2013). These heuristics may yield excellent or suboptimal results depending on the complexity of the situation (Simon, 1969, 1982). Behavioral explanations also recognize that decision framing can alter decisions (Kahneman et al., 1982), that situational factors such as time pressure and poverty consume scarce cognitive resources that can lead to poor decisions (Shah et al., 2012) and that decisions made in conditions of stress can be strongly conditioned by fear, anger, and other psychophysiological reactions (Lo and Repin, 2002; Rudolph and Repenning, 2002).

Are hoarding and phantom ordering rational, strategic responses to scarcity, or emotional, behavioral reactions triggered by stress? What may look like hoarding and phantom ordering could be rational responses to scarcity, particularly when there is uncertainty about final demand, supplies are subject to stochastic shocks and interruptions, capacity constraints may limit production, the consequences of shortages are high, storage and order cancellation costs are low, and multiple customers compete for limited supplies (Lee et al., 1997; Cachon and Lariviere, 1999; Armony and Plambeck, 2005).

Alternatively, scarcity may cause stress, anxiety, fear or panic, leading people to build their private stocks or place orders for more than they need even when it is not rational to do so. To illustrate, gasoline supply disruptions have sometimes caused retail service stations to run out, leading to “Sorry—No Gas” signs; episodes include the 1979 gas crisis in the US, transport strikes in Europe in 2000, and the aftermath of Superstorm Sandy on the east coast of the US in 2012. In each case, gas shortages led people, including many with nearly full tanks, to queue for fuel, often for hours; the long lines themselves then increased the perception of shortage in a positive feedback. For example, after Superstorm Sandy, “... drivers waited in lines that ran hundreds of vehicles deep, requiring state troopers and local police to protect against exploding tempers.

... The lines themselves only exacerbated the problem; reports in the local media provoked drivers to buy gasoline before stations ran out. Some spent what fuel they had searching for more and could be seen pushing vehicles toward relief.

‘I just want to have it, because you don’t know how long this is going to last,’ said Richard Bianchi, waiting in the half-mile line at the Sunoco in Union [New Jersey] with a tank that was three-quarters full. ‘People are panicking,’ said Jimmy Qawasmi, the owner of a Mobil in the Westchester County town of Mamaroneck.”

Of course any situation may involve a mix of strategic, rational action and behavioral, emotional responses, and it is difficult to disentangle the contribution of each in naturalistic settings. Here we explore the extent to which hoarding and phantom ordering are behavioral phenomena through an experimental study using the Beer Distribution Game (BDG). We use the data collected and reported by Croson, Donohue, Katok and Sterman (2014), hereafter CDKS. The CDKS experiment followed standard protocols in experimental economics, including financial reward for each participant proportional to supply chain profit. Experimental conditions eliminate all operational causes for supply chain instability generally and for hoarding and phantom ordering specifically. There are no random shocks to demand or supply, no capacity constraints on

---

production, and no price variations. Each supplier has only one customer so there is no horizontal competition to trigger short-age gaming. No order cancellations are permitted. Further, in the CDKS implementation of the BDG, there is no uncertainty about final demand; indeed, customer demand is constant at all times at four cases per week, and this fact is publicly announced in advance to all participants. Under these conditions rational agents would never create a bullwhip effect, generate oscillations, hoard inventory or place phantom orders.

Nevertheless, CDKS finds a significant bullwhip effect and oscillations in nearly all cases. Even more surprising, despite common knowledge that customer demand is constant at four cases per week, 22% of participants order hundreds of cases at a time or more—more than 25 times larger than their common knowledge of the constant demand—even though doing so destabilizes the supply chain, raises costs and reduces their financial reward. CDKS treat the extreme cases as outliers, and, although CDKS show that their results hold whether the outliers are included or excluded, they do not attempt to explain how such behavior arose.

To explore the role of behavioral processes in the BDG, we estimate various decision rules for participant orders. Our decision rules allow for endogenous hoarding and phantom ordering by relaxing the assumption in prior work that desired on-hand and on-order stocks are constant. We find 36% of participants engaged in statistically significant hoarding or phantom ordering.

Why would people engage in hoarding and phantom ordering even when it is irrational to do so? Mammals including humans evolved in environments characterized by variability, uncertainty, and competition for scarce resources. Hoarding is often an adaptive behavior in such circumstances, and is observed in many species. Neuroanatomical and brain imaging studies suggest that the propensity to hoard emerged early in mammalian evolution, and that the brain centers responsible for hoarding are distinct from the loci of economic decision-making. The Diagnostic and Statistical Manual of Mental Disorders (DSM-5) now recognizes hoarding as a distinct psychiatric disorder affecting 2-5% of the population (American Psychiatric Association, 2013). Brain imaging studies of such individuals find less activity in areas of the brain associated with decision making and emotional control, suggesting that the limbic system, where emotional responses and the impulse to hoard arise, is able to dominate behavior (Saxena et al., 2004; Anderson et al., 2004).

Does the prevalence of hoarding disorder in the general population mean hoarding and phantom ordering are inevitable, that managers must resign themselves to the instability and costs they create? If the propensity to engage in hoarding or phantom ordering is an attribute of individuals, then the incidence of these behaviors among participants in the experiment should be distributed randomly. If this dispositional hypothesis is correct, supply chain managers might seek to screen out job candidates who suffer from hoarding disorder or are predisposed to it. Alternatively, everyone may be susceptible to hoarding and phantom ordering to some degree, with these behaviors triggered by stressors in the situation such as unexpected delivery problems or demand spikes. If this situational hypothesis is correct the incidence of hoarding and phantom ordering will not be randomly distributed; rather, participants who experienced stressors such as large, unexpected demand surges or poor supplier performance will be more likely to engage in hoarding or place phantom orders than those who did not experience such stressors. If stress can cause irrational hoarding to overwhelm rational decision making even in normal individuals, screening won’t work; instead firms should seek to redesign their supply chains to minimize stressors. We test the dispositional and situational hypotheses in the CDKS data. The dispositional hypothesis is strongly rejected, while the situational hypothesis is supported: the incidence of hoarding and phantom ordering is much greater among participants who experienced stresses including order spikes, backlogs, and poor supplier performance.

Why does it matter whether hoarding and phantom ordering are strategic or behavioral phenomena? Strategic agents will engage in these behaviors when it is rational to do so, and will not if the system is redesigned so that hoarding and phantom ordering are no longer optimal. In contrast, if hoarding and phantom orders arise as the result of bounded rationality and emotional arousal, these behaviors can be triggered by scarcity even among normal individuals and even when it is irrational to do so, and effective interventions require more than an optimal policy.

We next review the literature on hoarding and phantom ordering, showing how positive feedback causes scarcity to be self-reinforcing, and contrast the rational and behavioral theories of the phenomenon. Section 3 describes the experiment and results. In section 4 we estimate various ordering decision rules, relaxing the restrictive assumptions of prior work that desired on-hand and on-order stocks are constant. In section 5 we test the dispositional and situational hypotheses. In the discussion we explore the implications and review results from neuroscience and psychiatry that bear on the etiology and evolutionary origins of hoarding. We close with implications for supply chain design and future research in behavioral operations management.

2. Hoarding, phantom orders and positive feedback

Phantom ordering has long been recognized. Mitchell (1924) provided an early description of the process and its role in supply chain instability. He described a situation in which

“Retailers 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 per cent. . . . They hope to be able to give full service next season, by which time, no doubt, these unexplainable conditions will have been remedied. However, retailers, having been disappointed in deliveries . . . are not going to be caught that way again. . . . Next season, if they want 100 units of an article, they order 100, so as to be sure, each, of getting the 90 in the pro rata share delivered. . . . Furthermore, to make doubly sure, each merchant spreads his orders over more sources of supply. . . . The false demand is passed back, stage by stage, along the channels of production. . . . What, in turn, is the natural result? . . . Eventually the streams of production are not only enlarged but overenlarged. There comes a time when the ultimate sources of supply fill nearly all the orders of their customers. The latter are surprised to find their orders filled promptly and fully. . . . There is no longer a shortage. . . . Producers in the second stages also fill their orders promptly and fully, thus surprising their customers in turn. . . . And so on down to the retailers. . . . Production has come to exceed the real demand.”

(Mitchell 1924: 645–647).

Hoarding and phantom ordering are common in many industries (Goncalves, 2003). Phantom orders are particularly prevalent in industries with long and variable lead times such as chemicals, pulp and paper, semiconductors, and high-tech (see e.g., Burton et al., 2005 for an example in the chemicals industry and Shi, 2002 for an example in high technology). Hoarding, placing multiple orders, and panic buying are also frequently seen for hot consumer products such as toys, games, and electronics (e.g. Cabbage Patch dolls, X-Box, iphones). Panic buying and hoarding of food, fuel, and medicines are common during wartime and periods of civil unrest, in regions where goods are chronically scarce (e.g., Russia and eastern Europe under communism), after disasters (e.g., superstorms Sandy and Haiyan) and in times of fear and uncertainty (e.g., the 1979 US gasoline crisis, the run-up to Y2K). The 2001 anthrax attacks in the US led to runs on the antibiotic Ciprofloxacin (Cipro).
Rationing of flu vaccine in the fall of 2004 led to panic buying, followed within months by surplus stocks (US House Committee on Government Reform, 2005), and fear of avian flu triggered hoarding of the antiviral drug oseltamivir (Tamiflu) (Pollack 2005).

Fig. 2 shows the feedback structure leading to hoarding and phantom ordering (Sterman 2000; Ch. 17–18, provides more detail; Forrester (1961) formulated the first dynamic supply chain model with endogenous desired inventory coverage and desired supply lines of on-order inventory). The figure shows the interaction between a customer, distribution channel (e.g., a retailer), and supplier. To maintain adequate service levels for its customers, the distribution channel maintains an inventory of product. The channel also maintains a supply line of orders with the supplier (the channel order backlog). Channel managers place orders with their supplier to (1) replenish units sold to customers, (2) close any gap between their desired and actual on-hand inventory level, and (3) close any gap between the desired and actual supply line of goods on order with the supplier. These actions form several negative feedbacks.

First, if channel inventory falls below the target level, channel managers (or their inventory management software) order more from the supplier. The supplier increases shipments to the channel, though it takes time to do so (the delay is shown in the link between the supplier’s desired shipments and actual shipments). Deliveries to the channel raise on-hand inventories, closing the negative (balancing) Inventory Control loop, B1. Second, distribution channel managers adjust orders to maintain the supply line of unfilled orders at the level sufficient to ensure adequate and timely deliveries. For example, if the channel needs to receive 1000 units from the supplier each day, and it takes the supplier 10 days to fill orders, then, by Little’s Law, the channel must maintain, in equilibrium, a supply line of 10,000 units on order. If, through process improvement, the supplier can cut the normal delivery time to 5 days, then the channel need maintain a supply line of only 5000 units. Channel managers, or their software, would cut orders until the supply line fell from 10,000 to 5000, closing the Supply-Line Control loop, B2. Third, an increase in its backlog leads the supplier to increase shipments, reducing the backlog of outstanding orders and closing the negative Availability loop, B3.

Under normal circumstances, the supplier is able to fill orders fully within the normal lead-time. Channel managers are content with small safety stocks (safety stock would only be required to buffer against uncertainty in customer demand, following standard inventory theory). The dynamics of the system are dominated by the three negative feedbacks described above. Now, however, imagine that the supplier’s ability to ship falls short of requirements. The cause could be the introduction of a surprisingly popular product, an uptick in the economy, capacity constraints, transportation bottlenecks, strikes or natural disasters. As shipments fall below desired levels, the supplier’s lead time rises. To ensure a steady supply of deliveries, channel managers must now order farther ahead. In the example above, if the supplier’s delivery time rises from 10 to 15 days, channel managers must (in equilibrium) keep 15,000 units on order. Further, as the probability of missed deliveries and partial allocations rises, channel managers may seek to have more than 15,000 units on order to keep the chance that deliveries fall short of requirements to an acceptable level. Orders rise further, causing delivery times to rise still more while allocations fall and delivery reliability drops, closing the positive (reinforcing) Phantom Orders feedback, R1. Furthermore, facing longer delivery delays and less reliable deliveries, channel managers may increase their target safety stocks. To do so they must order more from their suppliers, further lengthening delivery delays and reducing supplier reliability, triggering still larger target safety stocks and closing the positive Hoarding feedback, R2. Contrary to the normal functioning of markets, scarcity increases demand.

Fig. 3 shows the typical dynamics. Final demand follows a standard lifecycle, rising after product introduction, peaking, and then trailing off as the product ages. Rising sales deplete distribution channel inventory. To restore inventories to desired levels the channel must raise orders placed with suppliers above final demand. If supplier capacity is constrained or slow to respond, supplier lead times rise and the product may be placed on allocation. Distribution channel managers respond by placing phantom orders and increasing their safety stock targets, boosting orders still further above final sales. As Mitchell describes, each firm orders more than they want to compensate for shrinking allocations, further increasing demand.

Note that the dynamics above follow from fundamental queuing theory, e.g., the fact that, by Little’s Law, equilibrium on-order quantities must equal the product of desired deliveries and the delivery delay. A large literature explores more subtle effects, such as whether low inventory lowers sales by increasing consumers’ search costs or raises sales by signaling scarcity. Cachon et al. (2013) provide a review of the literature and an empirical test of these effects in the US automobile industry, finding that lower inventory variety cuts retail sales but that lower inventory of the same vehicles increases sales (a scarcity effect).

Please cite this article in press as: Sterman, J.D., Dogan, G., “I’m not hoarding, I’m just stocking up before the hoarders get here.”. Behavioral causes of phantom ordering in supply chains. J. Operations Manage. (2015). http://dx.doi.org/10.1016/j.jom.2015.07.002
ordering or shortage gaming is rational when there is horizontal competition among customers. However, if shortages trigger emotional arousal, customers may order more and seek to build safety stocks out of panic and fear rather than strategic consideration of how others will act. Either way, rising orders lead to still longer delivery times and still greater orders, making it difficult to determine the role of rational versus behavioral factors. Laboratory experiments, on the other hand, can be designed to eliminate all strategic incentives for hoarding and phantom ordering. In such a setting, rational agents would never hoard or place phantom orders, while these behaviors may still be observed if they arise from behavioral factors and emotional arousal.

3. The beer distribution game with constant, known demand

The Beer Distribution Game (BDG) is widely used in operations management as both a teaching and research tool. It has been used to understand the sources of the bullwhip effect, and the impact of POS data sharing, inventory visibility, and learning and communication on supply chain performance (Sterman, 1989a; Chen et al., 2000; Croson and Donohue, 2003; Steckel et al., 2004; Croson and Donohue, 2005, 2006; Croson et al., 2014; Wu and Katok, 2006). These studies show that participants endogenously generate substantial supply chain instability, including amplification in order variability from final demand to production, oscillations, and phase lag (the peaks of orders tend to be later as one moves from final demand to production); these phenomena are common in many industries (e.g., Sterman, 2000; Bray and Mendelson, 2012).

Lee et al. (1997) identify four operational causes of the bullwhip effect, including order batching due to quantity discounts, forward ordering based on price expectations (ordering more when prices are expected to rise and less when they are expected to fall), shortage gaming induced by horizontal competition as customers compete for limited supply, and demand forecast uncertainty. The standard protocol for the BDG (Sterman, 1989a) eliminates the first three. There are no quantity discounts or scale economies, so there is no incentive for order batching. Prices are fixed and demand is exogenous, so there is no incentive for forward ordering. There are no random shocks and no capacity constraints. Most important in the context of phantom ordering, each supplier serves only one customer and each customer orders from only one supplier. Thus there is no incentive to outcompete others for greater allocations by over-ordering. However, in most prior experiments customer demand was unknown to the participants and/or nonstationary. In such situations, amplification may arise from demand forecast updating. Croson and Donohue (2003, 2006) eliminated this possibility by informing all participants that demand is drawn from a specified stationary uniform distribution. Amplification and instability remained. Still, one might object that although subjects knew the distribution of demand, the realizations of customer orders were unknown (ex ante).

CDKS (2014) go further by making demand completely constant and known to all participants. They implemented the BDG using a web-based interface to prevent accounting or measurement errors and to control communication among participants. The experiments followed standard protocols in experimental economics. As in the standard game (Sterman, 1989a), holding costs were $0.50/case/week and backlog costs were $1.00/case/week.

Unlike prior studies, customer demand was constant. Further, this fact was publicly announced, inducing common knowledge among all participants. The realization of customer demand (four cases/week) was also displayed on each participant's screen every period, serving as an additional reminder that demand was constant. Prior to the start of each session participants took a quiz.
testing whether they understood the cost structure and rules of the game and that demand was constant at four cases/week; play did not commence until all answered correctly. After the session, but before receiving payment, participants were asked to complete a questionnaire asking them to describe their strategy and their beliefs about their teammates’ strategies. Participants were randomly assigned to experimental treatments and roles in the supply chain. CDKS (2014) provide full documentation.

With constant demand and common knowledge of that fact all operational causes of supply chain instability are eliminated, the amplification and oscillations seen in prior experiments should disappear. Furthermore, under these conditions it is never rational to hoard or engage in phantom ordering. Participants who order more than needed to satisfy the known, constant demand would necessarily induce a backlog upstream, raising costs and lowering their payoff. CDKS examine a variety of experimental treatments. In all cases, they initialize each supply chain in flow equilibrium with four cases in each shipping, production, and order processing stage. Therefore, if all participants order four cases/week, inventories would remain constant. In the base condition, on-hand inventories are initialized at the cost-minimizing equilibrium level of zero. The optimal strategy in that case is clearly for all participants to order four cases per week throughout, ensuring that inventories remain at the cost-minimizing level of zero at all times.

Nevertheless, nearly all the supply chains exhibit oscillation and amplification (increasing order variability from retailer to factory), and the bullwhip effect is statistically significant in all treatment conditions. Fig. 4 shows a typical team.

Two questions must be answered to understand these remarkable results. First, why do participants first depart from equilibrium by ordering a quantity other than the optimal, equilibrium value of four cases per week? Second, once the system is perturbed, why do participants order in such a way as to cause persistent amplification and oscillation, rather than returning to equilibrium in a swift, stable and cost-minimizing fashion?

CDKS argue that the answer to the first question is coordination risk. Though all participants have common knowledge of customer demand, some may not believe that their teammates will use the optimal strategy. To buffer themselves against the risk that others will not order optimally, these participants attempt to build up “coordination stocks” of inventory. Doing so requires that they order more than the optimal order quantity of four cases, as seen in Fig. 4, where all but the retailer immediately order more than four. Participants’ responses to the post-game questionnaire suggest many did not trust their teammates to play optimally and sought to increase their inventories as a hedge against coordination risk (see the supplement).

To answer the second question, CDKS estimated the ordering decision rule first tested in Sterman (1989a). Consistent with that study, they found that most subjects significantly underweighted the supply line of on-order inventory. Underweighting the supply line causes the system to become underdamped, and often locally unstable, generating oscillations in response to perturbations (Mosekilde and Laugesen, 2007; Thomsen et al., 1992; Larsen et al., 1999). Failure to account for the supply line of unfilled orders, and time delays generally, has been observed in a wide range of experimental studies beyond the beer game, including single-person settings with no strategic interactions (Diehl and Sterman, 1995), and many other dynamic decision making tasks (Sterman, 1989b; Breher, 1992; Kleinmuntz, 1993; Paich and Sterman, 1993; Dörner, 1996; Kampmann and Sterman, 2014). Failure to attend to delays and the supply line has also been shown to play important roles in industries from, as Sterman (2000, Ch. 20) notes, “aircraft to zinc” (see also Meadows, 1970; Randers and Göltuke, 2007; Pierson and Sterman, 2013 for examples). The tendency to ignore or underestimate time delays is robust to context, the availability of information about the supply line, incentives and opportunities for learning.

CDKS also identify a number of outliers who placed very large orders. Maximum orders were 100 cases/week or more for 22% of the participants in all treatments with humans in each role, an astonishing 25 times greater than the publicly known, constant customer demand of four cases/week.4 Fig. 5 shows the most extreme case. The key participant is the wholesaler, who orders 30,000 cases in week 12 even though retailer orders average four cases/week and never exceed 36. The retailer, apparently seeking to build a safety stock, begins by ordering 16 cases, forcing the wholesaler into a backlog. Initially the wholesaler responds modestly to the retailer’s unexpected behavior, increasing orders to 8 cases in week 3. But the wholesaler soon finds that the distributor is not able to deliver at that rate: production has not yet increased, so deliveries continue at 4 cases/week. Worse, the distributor actually cuts orders, so in week 7 the wholesaler unexpectedly receives only 2 cases. The wholesaler responds to the drop in supplier deliveries with still higher orders, causing the distributor’s service level to plummet. By week 9 the wholesaler receives less than 0.2% of the supply line on order with the distributor, and responds by ordering 1000 cases.

5 CDKS examine experimental treatments including (i) publicly providing participants with the optimal decision rule (indicating common knowledge of optimal strategy), (ii) replacing all participants but one with automated agents playing the optimal decision rule (eliminating uncertainty about whether others will play optimally), and (iii) providing initial inventory (buffering participants against the risk others may not behave optimally). Amplification and oscillations in orders and inventories, though moderated by the treatments, remained statistically significant in all conditions. The tendency to ignore the supply line was not moderated by the experimental treatments.

6 Note that a participant who places a large order could still be rational: optimal play given an incoming order of, say, 200 more than the equilibrium rate is to order 200 and then return to the equilibrium order (see note 7). Hence it is possible that not all of the 22% who ordered 100c/wk or more at least once deviated from rational play. However, CDKS show: (1) more than 70% of participants spontaneously deviated from the optimal order of 4c/wk; (2) there is statistically significant order amplification, which is suboptimal; and (3) the vast majority of participants underweight the supply line, leading to suboptimal over-accumulation of inventory and oscillations. Note also that CDKS show that their results, including supply-line underweighting, hold with outliers included or excluded.
forcing the distributor’s service level to fall further, leading the wholesaler to order still more. By week 12 the wholesaler receives less than 0.014% of the supply line on order. If that low service level were expected to persist it would be necessary to order more than 28,000 cases to receive 4. In fact, the wholesaler orders 30,000. In week 13 the wholesaler apparently realizes—finally—that the supply line is enormous and slashes orders. Unable to cancel orders, the wholesaler accumulates surplus stock—more than 50,000 cases.

4. Modeling hoarding and phantom ordering

To test the hypothesis that extreme orders arise from hoarding and phantom ordering, we now estimate formal decision rules to model participant ordering behavior. We use the CDKS dataset, consisting of all the participants reported in CDKS (2014), including outliers. We also include participants from two treatments CDKS did not report in the published paper, a total of 240 participants.\(^5\)

We omit factories from our analysis (60 participants) because, with no capacity constraint, factories always receive their entire order with a constant delivery time, which means their propensity to phantom order cannot be estimated. We also omit nine participants (3.75%) who remained in equilibrium throughout the game. Without any variability it is not possible to estimate any decision rule for these individuals. The resulting dataset consists of 171 participants.

CDKS detail participant demographics and summary statistics for

\(^5\) These were (i) all human teams with initial on-hand inventories of 12 cases and common knowledge of the optimal decision rule, and (ii) teams with three optimal automated agents and initial on-hand inventories of 12 cases; the human participant was provided with the optimal ordering rule in addition to knowledge of customer demand.
the participants they report; the supplement to this paper contains all the models and data needed to replicate our analysis.

We begin with the decision rule used by Sterman (1989a) and used in subsequent research. The rule is based on the anchoring and adjustment heuristic (Tversky and Kahneman, 1974). Participants are assumed to anchor on the quantity they expect their customer to order from them (D^\ast, expected demand), then adjust this quantity so as to (1) reduce the discrepancy between the desired and actual stock (the stock adjustment, A^S) and (2) reduce the discrepancy between the desired and actual supply line of goods on order (the supply line adjustment A^SL). Since cancellations are not permitted, orders must be nonnegative:

\[ O_t = \max \{ 0, D_t^f + A_t^S + A_t^SL \} \]  

(1)

Sterman (1989a) and CDKS (2014) modeled expected demand using exponential smoothing. In the CDKS experiment, however, participants are publicly informed that customer orders are constant at four cases/week at all times. Some participants appeared to use their common knowledge of customer demand in placing their orders while others adjusted their estimates of incoming demand as it varied. To capture these possibilities, we formulate expected demand as a weighted average of the commonly known constant customer demand, D^\ast, and the participant’s belief about the orders they expect to receive from their customer, D^f, which, as in Sterman (1989a), is formed by exponential smoothing of the orders the participant actually receives, D:

\[ D_t^f = \psi D_t^f + (1 - \psi) D_{t-1}^f \]  

(2)

\[ D_t^p = \theta D_{t-1} + (1 - \theta) D_{t-1}^p \]  

(3)

where \( \psi \) is the sensitivity of expected demand to customer orders and \( \theta \) is the smoothing parameter on actual incoming orders. If \( \psi = 1, D^f \) always equals the commonly known constant customer order rate of four cases/week. Values \( 0 < \psi < 1 \) mean expected orders respond to some extent to smoothing incoming orders.\(^6\)

The stock adjustment is assumed to be linear in the discrepancy between desired inventory, S^*, and net inventory, S:

\[ A_t^S = \alpha_S (S_t^* - S_t) \]  

(4)

where \( \alpha_S \) is the fraction of the inventory discrepancy ordered each period. The supply line adjustment is also linear in the discrepancy between the desired supply line, SL^*, and the actual supply line, SL:

\[ A_t^SL = \alpha_{SL} (SL_t^* - SL_t) \]  

(5)

where \( \alpha_{SL} \) is the fraction of supply line discrepancy ordered each period. Combining terms:

\[ O_t = \max \{ 0, D_t^f + \alpha_S (S_t^* - S_t) + \alpha_{SL} (SL_t^* - SL_t) \} \]  

(6)

To estimate the parameters, Sterman (1989a) defines \( \beta = \alpha_{SL}/\alpha_S \), so estimated orders, O_t, become

\[ O_t = \max \{ 0, D_t^f + \alpha_S (S_t^* - S_t) + \beta (SL_t^* - SL_t) \} \]  

(7)

where \( \beta \) is the fraction of the supply line the participant takes into account, \( 0 < \beta < 1 \). Sterman (1989a) and CDKS further assume that the desired stock and desired supply line are constant. In this case \( S^* \) and \( SL^* \) can be independently estimated. However, defining effective total desired on-hand and on-order stock \( S^* = S^* + \beta SL^* \), and assuming an additive error term, the system to be estimated, denoted Model 0, becomes:

\[ O_t = \max \{ 0, D_t^f + \alpha_S (S_t^* - S_t - \beta SL_t^*) + \epsilon_t \} \]

\[ D_t^f = \psi D_t^f + (1 - \psi) D_{t-1}^f \]

\[ D_t^p = \theta D_{t-1} + (1 - \theta) D_{t-1}^p \]  

subject to \( 0 \leq \psi, \theta, \alpha_S, \beta \leq 1; 0 \leq S^* \)

The optimal parameters are readily determined. In CDKS all participants have common knowledge that customer demand is constant and that there are no random events. Participants should expect incoming orders will equal the known, constant customer order rate of four cases/week. Therefore, \( \psi = 1 \) (and \( \theta \) is irrelevant).\(^7\)

Similarly, with common knowledge of constant customer demand, no stochastic effects and the assumption of rational behavior, there is no need for any safety stocks, so \( S^* = 0 \). The optimal supply line, SL^*, is readily determined. By Little’s Law, a participant who seeks to receive goods at the rate of \( R^* \) cases/week must maintain, in equilibrium, a supply line equal to

\[ SL^* = \lambda^S R^* \]  

(9)

where \( \lambda^S \) is the expected lag in receiving orders. Since customer demand is constant and that fact is common knowledge, the equilibrium rate at which each participant seeks to receive deliveries from their supplier \( R^* = 4 \) cases/week. In the BDG, the normal, equilibrium delivery lag, \( \lambda^* \), is 3 weeks for the factory and 4 weeks for all others. Hence, given the common knowledge and assuming optimal play (and hence that \( \lambda^* = \lambda^S \)), SL^* = 12 cases for factories and 16 for all others. The optimal value of \( \beta \) is one: participants should fully account for the supply line of orders placed but not yet received (Sterman, 1989a). The optimal value of \( \alpha_S \) is also one: since there are no adjustment costs, participants should strive to eliminate any inventory shortfall or surplus as quickly as possible so as to minimize holding and backlog costs.

The optimal decision rule, with constant expected demand and \( \alpha_S = \beta = 1 \), reduces to the familiar order-up-to rule in which participants order, each period, the entire shortfall between desired and actual inventory, where actual inventory includes both on-hand and on-order stocks (Clark and Scarf, 1960). If inventories differ from the optimal value, rational agents will alter their orders temporally until sufficient orders to correct any inventory discrepancy are in the supply line, then return to the equilibrium order quantity. There would be no oscillation.

Following CDKS we use nonlinear least squares to estimate the decision rule in Eq. (8), denoted Model 0, and bootstrapping to estimate the confidence intervals for each parameter (the supplement provides details). Table 1 shows the mean and median estimates for Model 0, which assumes constant desired on-hand and on-order stocks.

Most of the estimated parameters are far from optimal. The median estimate of \( \psi \) is 0.21, and the estimates of \( \psi \) are statis-
2. “Try to use the base-stock policy. Base-stock level equal to the lead time demand=12, since I am the manufacturer with 3 weeks delay. The incoming orders may force me to raise the inventory.” [Manufacturer]

3. “I tried to keep it so I had a 2 week inventory available to me at all times.” [Retailer]

4. “Because I was aiming for zero inventory, I went negative a couple of times… I tried to keep a little inventory later to anticipate spikes….” [Manufacturer]

The post-session questionnaire also asked participants “What caused your inventory/backlog to behave as it did?” Many cited unreliable deliveries from their supplier:

5. “The supplier not fulfilling the order in the time you wanted it to” [Retailer]

6. “The lack of steady shipments.” [Retailer]

7. “The distributor was not capable to deliver [sic] what I ordered.” [Wholesaler]

8. “Uncertainty of what [the] distributor [was] sending, uneven amounts.” [Wholesaler]

9. “I didn’t know the orders that I would receive or how much the manufacturer was producing.” [Distributor]

These responses suggest that some participants changed their desired inventory and desired supply line in response to unexpectedly large incoming orders or variations in the service levels and delivery time of their supplier. Formally, we hypothesize that some participants may have set their desired stock to maintain a certain number of weeks of coverage of expected incoming orders, as suggested by participant comments 1-4 above:

\[ S_{t}^* = \gamma dR_t \]

where desired coverage, \( \gamma \), is a parameter to be estimated.

Turning to the desired on-order stock, recall that Little’s Law requires that (in equilibrium) the supply line be set to enable deliveries to be equal to the desired delivery rate, \( R_t \), given the expected delivery delay, \( \lambda^D \) (Eq. (9)). Model 0 assumes \( R_t \) and \( \lambda^D \) are constant. We now allow for the possibility that both may vary as participants experience unreliable supplier delivery performance (as suggested by comments 5-9 above):

\[ SL_{t+1} = \lambda^D \frac{S_{t+1}}{R_t} \]

We formulate the expected delivery delay as a linear function of the participant’s belief about the current delivery delay, denoted the perceived delay, \( \lambda^P \), but constrained to be less than some maximum, \( \lambda^M \):

\[ \lambda^D_t = \min\{\lambda^M, \kappa + \omega^L_{t-1} \} \]

where \( \kappa, \omega \), and \( \lambda^M \) are parameters to be estimated. Rational agents would set \( \kappa \) to the equilibrium delivery delay, \( \kappa = \lambda^D \), and would never update their estimate of the delivery delay (\( \omega = 0 \)) because they know that any change in supplier lead times is temporary. But to the extent \( \omega > 0 \), the belief that supplier lead times have risen increases \( S_{t+1}^* \), causing phantom orders.

How might participants estimate the current delivery delay, \( \lambda^D \)? By Little’s Law, the steady-state delay is given by the ratio of the supply line to deliveries received, \( SL/R \). However, participants know

| Table 1 Summary of estimation results for Model 0 (constant total desired inventory).^a |
|-----------------|-----------------|-----------------|
| Model 0         | Parameter      | Mean   | Median | n   |
|                 | \( \psi \)      | 0.32   | 0.21   | 108  |
|                 | \( \theta \)    | 0.33   | 0.25   | 107  |
|                 | \( \alpha_1 \)  | 0.32   | 0.22   | 171  |
|                 | \( \beta \)     | 0.34   | 0.34   | 161  |
|                 | \( S \)         | 132.4  | 7.44   | 161  |
|                 | \( R_t^* \)     | 0.46   | 0.44   | 171  |
|                  | RMSE            | 89.29  | 2.63   | 171  |
| Subjects        | 171             |        |        |      |

^a The number of estimates, \( n, \psi, \theta, \beta \) and \( S \) is less than the number of participants (N = 171) because it is not possible to estimate \( \psi \) or \( \theta \) if incoming orders are constant (including all retailers); not possible to estimate \( \theta \) if \( \psi = 1 \), and not possible to estimate \( \beta \) if \( \alpha_1 = 0 \).
that the delivery delay can never be less than the normal value \( \lambda^* \), even if the supply line falls to zero (which can occur if the participant orders zero for a sufficiently long period, a situation resulting from excess inventory). Similarly, participants may receive nothing from their supplier, but we assume that their estimate of the delivery delay remains finite by constraining expected deliveries to the maximum of actual deliveries, \( R_t \), and 1 case/week:

\[
\lambda^P_t = \max(\lambda^*, S_t / \max(R_t, 1))
\]  

(13)

The questionnaire responses do not reveal how participants determine \( \lambda^P \), the desired rate at which they seek to receive goods from their supplier. Following Sterman (2000, Ch. 17), we consider two alternatives. In the first, denoted Model 1, participants seek to receive goods at the rate they expect their customers will place orders:

Model 1: \( R_t^1 = D_t^e \)  

(14a)

In the second, denoted Model 2, participants desire to receive goods at the rate needed to fill expected customer orders and eliminate any inventory gap they have:

Model 2: \( R_t^2 = \max(0, D_t^e + \alpha_5(S_t - S)) \)  

(14b)

where the max function ensures that \( R_t^2 \) remains nonnegative even if there is surplus inventory.

To summarize, prior studies (Sterman, 1989a; Diehl and Sterman, 1995; Croson and Donohue, 2006; CDK5, 2014) estimate decision rules that rule out hoarding and phantom ordering because desired on-hand and on-order inventories are constant by assumption. We relax these restrictive assumptions to allow for the possibility of hoarding (increases in desired on-hand inventory) and phantom orders (increases in desired on-order inventory), both arising endogenously in response to demand variations or poor supplier delivery performance.

For the desired on-hand stock, we assume participants seek a constant desired coverage of expected orders. For the desired on-order stock, \( \lambda^P_sR^s \), we consider two possibilities for the desired acquisition rate, \( R^s \). We estimate both alternative models for each subject by nonlinear least squares, using bootstrapping to estimate the confidence intervals. We define the best model for each participant to be the model that provides the lowest RMSE among Models 0–2.

Table 2 summarizes the results. The models allowing hoarding and phantom ordering have lower RMSE than Model 0 for 119 (69.6%) of the participants. Model 1, in which \( R^s = D^e \), fits best for 62 (36.3%) of the participants, and Model 2, in which \( R^s = \max(0, D^e + \alpha_5(S^* - S)) \), fits best for 57 (33.3%). Compared to Model 0, the best alternative model lowers the median RMSE modestly (4.3%), but cuts the mean RMSE by 39%; the alternative models provide the greatest improvement for the outliers—those for whom Model 0 offers the least explanatory power.

Fig. 6 shows the distributions of the estimated parameters capturing hoarding and phantom ordering (\( \gamma \) and \( \omega \)) for the 119 participants for whom Model 1 or 2 is best. Most do not seek large safety stocks, indicating little hoarding, with \( \gamma < 0.5 \) weeks for 51.3% of these participants. However, the distribution has a long tail (note the log-log scale for \( \gamma \)). Nearly one-quarter, 23.9%, seek more than two weeks of coverage, and 7.7% seek more than 16 weeks of coverage. The parameter \( \omega \) indicates the extent of lead-time updating in the formulation for the desired supply line. If \( \omega \approx 0 \), participants are not sensitive to the current delivery delay, indicating little or no propensity to place phantom orders. If \( \omega \approx 1 \), participants adjust the desired supply line in proportion to the current delivery delay, inducing phantom orders as supplier service levels drop. Values of \( \omega > 1 \) indicate even more aggressive phantom ordering—such participants order even more than the current delivery delay would indicate. The distribution of \( \omega \) also shows a long tail. Among those for whom Model 1 or 2 is best, \( \omega < 0.5 \) for 31.6%. However, \( \omega \geq 1 \) for 65.8% of participants, and \( \omega \geq 5 \) for nearly one-quarter (24.8%). These participants aggressively update their lead time estimates as supplier service levels fall, triggering large phantom orders that further increase the delivery delay they experience. There is little evidence that participants limited their estimates of the expected delivery delay. Eq. (12) constrains the expected delivery delay to a maximum value, \( \lambda^M \), potentially preventing extremely large phantom orders by capping the expected delivery delay even if the actual delay becomes very large. Although the maximum delivery delay experienced by the participants varies substantially, with a mean (median) of 651 (30) weeks, \( \lambda^M \) is a binding constraint for only 11.8% of the participants for whom model 1 or 2 is best.

Estimation results suggest roughly 70% of the participants engaged in hoarding or phantom ordering to some degree, as determined by the model yielding the lowest RMSE. A more stringent test, however, counts only those subjects whose estimated values of \( \gamma \) or \( \omega \) are statistically significantly greater than the optimal values of zero. We find \( \gamma \) is statistically significantly greater than the optimal value of zero at \( p < 0.025 \) for 27 individuals (15.8% of all 171 participants). We denote these individuals the hoarders. We find \( \omega \) is statistically significantly greater than the optimal value of

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best-fit decision rules among alternative models. *</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Median</td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>( \psi )</td>
<td>0.61</td>
<td>0.71</td>
<td>42</td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.38</td>
<td>0.27</td>
<td>34</td>
</tr>
<tr>
<td>( \alpha_s )</td>
<td>0.33</td>
<td>0.22</td>
<td>62</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.22</td>
<td>0.07</td>
<td>60</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>3.64</td>
<td>0.67</td>
<td>60</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>21.98</td>
<td>21.34</td>
<td>60</td>
</tr>
<tr>
<td>( \omega )</td>
<td>5.07</td>
<td>0.57</td>
<td>62</td>
</tr>
<tr>
<td>RMSE</td>
<td>24.55</td>
<td>2.32</td>
<td>62</td>
</tr>
</tbody>
</table>

* The number of estimates, \( n \), of \( \psi, \theta, \beta, \gamma, \kappa, \) and \( \omega \) is less than the number of participants for whom the alternative models fit better than Model 0 (\( N = 119 \)) because it is not possible to estimate \( \psi \) or \( \theta \) if incoming orders are constant (including all retailers); not possible to estimate \( \theta \) if \( \psi = 1 \), and not possible to estimate \( \beta, \gamma, \kappa, \) and \( \omega \) if \( \alpha_s = 0 \). Additionally, for most participants, the best fit is achieved with values of the maximum allowed value of the perceived delivery delay, \( \lambda^M \) (Eq. (12)), that exceed the largest value of the actual delivery delay, \( \lambda^* \), indicating that the constraint on the allowed value of the perceived delivery delay is not binding.

---

Please cite this article in press as: Sterman, J.D., Dogan, G., “I'm not hoarding, I'm just stockpiling up before the hoarders get here.”. Behavioral causes of phantom ordering in supply chains. J. Operations Manage. (2015), http://dx.doi.org/10.1016/j.jom.2015.07.002
zero for 45 individuals (26.3% of all participants). We denote these participants the phantom orderers. A total of 61 individuals were either hoarding or phantom ordering (35.7% of all participants), and a total of 11 individuals (6.4% of all participants) were both hoarding and placing phantom orders.

Fig. 7 compares the actual and best-fit ordering decisions for several of the outlier participants. The examples shown include instances for which both Models 1 and 2 provide the best fit. The alternative models explain the most extreme participants well, including the wholesaler who ordered 30,000 cases in a single week despite the fact that incoming orders never exceeded 36 (participant 174; see Fig. 5). That participant does not appear to engage in hoarding ($\gamma = 0$), but is extremely sensitive to supplier delivery performance ($\omega = 4.22, \lambda^M = 3075$). Model 1 reduces the RMSE by 83.3% compared to Model 0.

The estimation results support the qualitative analysis in section 4: as the supplier delivered a smaller and smaller fraction of the wholesaler’s orders, the wholesaler compensated by ordering more and more. The results illustrate how the positive feedback of phantom ordering (Fig. 2) can dominate the negative feedbacks that should limit orders and keep inventory close to the optimal level.

The supply line adjustment term in the ordering decision rule (Eq. (7)) is

$$A^SL = \alpha_{SL} (SL^* - SL_t) = \alpha_{SL} (SL^* - SL_t)$$

In models 1 and 2, the desired supply line is $SL^\ast = \lambda^\ast t R_t' \ (\text{Eq. (11)}),$ yielding

$$A^SL = \alpha_{SL} \left( \lambda^\ast t R_t' - SL_t \right).$$

Substituting the formulation for $\lambda^\ast$ (Eq. (12)) into this expression yields

$$A^SL = \alpha_{SL} \left( \min \left( \lambda^M, \kappa + \omega \lambda^P \right) R_t' - SL_t \right).$$

Using the formulation for the perceived delivery delay, $\lambda^P$ (Eq. (13)), gives

$$A^SL = \alpha_{SL} \left( \min \left( \lambda^M, \kappa + \omega \max \left( \lambda^*, SL_t / \max (R_t, 1) \right) \right) R_t' - SL_t \right).$$

We can gain intuition into the formulation by recognizing that phantom ordering is likely to arise during periods when suppliers are not able to deliver fully, which means the perceived delivery delay (Eq. (13)) will exceed the normal delay, $\lambda^\ast$. Assuming further that (i) $R_t \geq 1$, which is true initially and should be true during periods of high demand, and (ii) the maximum delivery delay, $\lambda^M$, is not a binding constraint, which estimation results show is true for the large majority of cases, the supply line adjustment becomes:

$$A^SL = \alpha_{SL} \left( \kappa + \omega (SL_t / R_t) \right) R_t' - SL_t$$

Collecting terms,

$$A^SL = \alpha_{SL} \left( \kappa + SL_t \left( \alpha(R_t^*/R_t - 1) \right) \right)$$

In the original decision rule with constant $SL^\ast$ (Eq. (8)), the actual supply line $SL$ always enters $A^SL$ and therefore orders with a negative sign: the larger $SL$, the smaller orders will be, reflecting the negative (balancing) Supply Line Control feedback (loop B2 in Fig. 2). However, as seen in Eq. (20), when the desired supply line responds to supplier performance, the net effect of the supply line adjustment on orders can be positive or negative. Specifically,

$$\omega(R_t^*/R_t) < 1 \Rightarrow \partial ASL / \partial SL < 0 \Rightarrow \partial O / \partial SL < 0 \Rightarrow \text{net negative feedback};$$

$$\omega(R_t^*/R_t) > 1 \Rightarrow \partial ASL / \partial SL > 0 \Rightarrow \partial O / \partial SL > 0 \Rightarrow \text{net positive feedback}. $$

If $\omega(R_t^*/R_t) < 1$, the larger the supply line, the larger orders will be, further increasing the supply line. In that case, the positive Phantom Ordering feedback (R1 in Fig. 2) dominates the negative Supply Line Control loop, leading to runaway growth in orders—until deliveries rise enough relative to desired deliveries for $\omega(R_t^*/R_t)$ to fall below 1 or the maximum expected delivery delay $\lambda^M$ becomes a binding constraint. The larger the value of $\omega$, the smaller the increase in $R_t^*/R_t$ required to cross the tipping threshold where the system shifts from net negative to net positive feedback.

Starting from the initial equilibrium, $R_t$ will exceed $R$ if incoming orders rise, causing an increase in expected orders $DF$, or (in Model 2) causing an inventory shortfall even if $DF$ remains constant (Eq. (14b)). $R_t$ can also exceed $R$ if the supplier’s deliveries fall.

The estimated parameters for participant 174 are $\psi = 1$ (making $\theta$ irrelevant), $\alpha_{SL} = 0.90, \beta = 1, \gamma = 0, \kappa = 0.01, \omega = 4.22,$ and $\lambda^M = 3075$. Because $\psi = 1$, the increase in incoming orders participant 174 receives from the retailer (peaking at 36 cases/week; Fig. 5) is not passed into expected demand. Desired inventory coverage $\gamma = 0$: the participant sets $S^\ast$ to the optimal value of zero and there is no evidence of hoarding. Since Model 1 is best for this subject, $R_t = DF$, and given $\psi = 1$, simulated $R_t$ never varies from the optimal value of 4 cases/week. However, the distributor cuts orders from 4 to 2 cases/week, causing deliveries to the wholesaler to fall beginning in week 7 (Fig. 5). Poor supplier delivery performance causes the simulated ratio $R_t^*/R_t$ to rise above one. Participant 174 reacts to the unexpected drop in supplier deliveries by increasing $SL^\ast$. With $\omega = 4.22$, even small increases in $R_t^*/R_t$ lead to net positive feedback. Further, since this participant fully accounts for the supply line ($\beta = 1$) and orders nearly the entire inventory shortfall.
each period ($\alpha = 0.90$) the gain of the positive phantom ordering loop, $\alpha \beta \rho \omega (R^*/R - 1)$, is extremely high. Orders and the supply line grow exponentially until the cap on expected delivery delay, $\lambda^M$, is reached (week 12), breaking the positive phantom ordering feedback. The actual supply line then surpasses the inflated target, and in week 13 orders collapse. The estimated parameters for participant 174 suggest that the unexpected drop in supplier delivery performance, not the rise in incoming orders from the retailer, caused the huge surge in orders as the wholesaler became caught in the positive phantom ordering feedback.

7. Dispositional or situational causes of hoarding and phantom ordering?

Why do some participants hoard or place phantom orders while others do not? One possibility is that certain individuals have a high propensity to hoard or place phantom orders, while others do not; under this dispositional hypothesis hoarding and phantom ordering are attributes of individuals. In terms of the models estimated above, desired inventory coverage, $\gamma$, indicative of hoarding, and the sensitivity of orders to perceived delivery delay, $\omega$, which indicates phantom ordering, would be distributed randomly among participants based on their individual propensity to engage in these behaviors (though those prone to hoarding may also be prone to phantom ordering, so the estimates of $\gamma$ and $\omega$ might be positively correlated).

Alternatively, everyone may be susceptible to hoarding and phantom ordering to some degree, with these behaviors triggered by stressors in the situation such as unexpected demand spikes or delivery problems. Under this situational hypothesis the incidence of hoarding and phantom ordering will not be randomly distributed; rather, participants who experienced stressors such as large unexpected demand surges or poor supplier performance will be much more likely to engage in hoarding or place phantom orders. Formally, we seek to assess

$H_0$, the Dispositional hypothesis: Hoarding and phantom ordering are attributes of individuals, not the situation in which individuals find themselves. Specifically,
Hoarding (as indicated by the estimated values of desired inventory coverage, $\gamma$) will be randomly distributed among participants, and uncorrelated with variability in incoming orders or supplier delivery performance.

(ii) Phantom ordering (as indicated by the estimated values of the sensitivity to supplier deliveries, $\omega$) will be randomly distributed among participants, and uncorrelated with variability in incoming orders or supplier delivery performance.

$H_0$, the Situational hypothesis: Hoarding and phantom ordering are triggered by stressors in the situation. Specifically,

(i) Hoarding (as indicated by the estimated values of desired inventory coverage, $\gamma$) is more likely when participants experience large surges in incoming orders or poor supplier delivery performance (because either can cause costly backlogs).

(ii) Phantom ordering (as indicated by the estimated sensitivity to supplier deliveries, $\omega$) is more likely when participants experience poor supplier delivery performance, but not large surges in incoming orders (because large inventories are not needed despite variability in incoming orders if suppliers can deliver quickly and reliably).

Among participants exhibiting statistically significant hoarding, the mean of the standard deviation of incoming orders is 227 cases/week compared to a mean of 23 cases/week for the non-hoarders, a highly statistically significant difference (two-tailed Mann-Whitney test, $U = 956, z = 4.185, p = 0.000014$). Similarly, the maximum delivery delay each hoarder experienced averaged 1765 weeks compared to a mean of 442 weeks for the non-hoarders; the difference is also highly statistically significant (two-tailed Mann-Whitney test, $U = 730, z = 5.143, p = 1.36 \times 10^{-7}$). The results are consistent with the Situational Hypothesis $H_0$ (i), while the Dispositional hypothesis $H_0$ (ii) is strongly rejected: Hoarding behavior is strongly associated with either unexpectedly large jumps in incoming orders or unexpectedly large delays in supplier deliveries.

Turning to phantom ordering, theory, beginning with Mitchell (1923), posits that firms will order more than they desire and place orders with multiple suppliers when they believe their suppliers cannot deliver fully or on time. In contrast, as long as suppliers can deliver quickly and reliably, there is no motivation to place phantom orders even if incoming orders vary. The maximum delivery delay experienced by the phantom orders is statistically significantly larger than that experienced by those who did not engage in phantom ordering (two-tailed Mann-Whitney test, $U = 1723, z = 3.897, p = 0.00001$). The results are consistent with $H_0$ (i), while $H_0$ (ii) is rejected: Phantom ordering is strongly associated with unexpectedly large increases in supplier delivery delay.9

The evidence supports the hypothesis that situational factors strongly affect whether participants engaged in hoarding or phantom ordering. Those who experienced unexpected variability in incoming orders or delays in deliveries from suppliers were more likely to hoard than those who did not. Those who experienced unexpected problems in receiving what they ordered from their supplier were more likely to place phantom orders. The results are consistent with a large role for situational factors as triggers for the irrational hoarding and phantom ordering observed in participant behavior.10

8. Discussion, limitations and conclusions

Hoarding and phantom ordering contribute to supply chain instability and inefficiency yet remain common in many industries and markets. Hoarding and phantom ordering can be rational for individuals (Lee et al., 1997; Cachon and Lariviere, 1999; Armony and Plambeck, 2005), though costly and destabilizing for the economy and society. Although hoarding and phantom ordering can result from the rational behavior of strategic agents, prior research, extending back at least to John Stuart Mill (1848), Mitchell (1924), and Forrester (1961), shows how such behavior can result from bounded rationality and emotional reactions. Many accounts of related phenomena such as speculative bubbles stress the role of behavioral and emotional factors, from Mackey’s (1841) classic account of the “madness of crowds” to Merton’s (1948) “self-fulfilling prophecy” to Shiller’s (2001) “irrational exuberance.”

The Beer Distribution Game provides an experimental setting in which rational and behavioral contributions to supply chain instability can be distinguished. The CDKS (2014) implementation of the BDG eliminates all operational and institutional structures that might cause rational agents to generate a bullwhip effect, hoard or place phantom orders, including order batching (because there are no scale economies favoring large order quantities), forward buying (because customer demand is exogenous and prices are fixed), uncertainty about demand (because customer demand is constant and that information is common knowledge), and shortage gaming (because there is no horizontal competition). Under these conditions rational agents would never generate supply chain instability, yet nearly all the experimental supply chains exhibited a pronounced bullwhip. Even more surprising, more than a fifth of all participants placed orders more than 25 times larger than the known, constant demand. CDKS showed that their results held whether these outliers were included or excluded. However, they did not examine why so many participants generated astonishingly large orders.

Although it is never rational in the CDKS setting to engage in hoarding or phantom ordering, we hypothesize that signals of scarcity, such as large increases in incoming orders or inadequate deliveries from suppliers, triggered hoarding or phantom ordering among many of the outlier participants. To test for the presence of hoarding or phantom ordering we estimated decision rules for participant orders, relaxing the assumption of prior studies that desired inventory and desired supply line levels are constant (e.g., Sterman, 1989a; Diehl and Sterman, 1995; Croson and Donohue, 2006; Croson et al., 2014). The alternative models allow for the possibility that people increase their desired inventory coverage (hoard) as they experience scarcity or increase the desired supply line of unfilled orders (phantom ordering) to compensate for changes in their supplier’s lead time. While most participants do not hoard or place phantom orders, we find a total of 35.7% of all participants exhibit statistically significant hoarding (15.8%), phantom ordering (26.3%), or both (6.4%). The alternative models significantly improve the explanatory power of the ordering decision rules for these subjects. The close match between the estimated and actual ordering behavior for these extreme participants suggests that their decisions did not reflect low motivation, weak incentives or poor understanding of the experiment, but rather placed by hoarders induce longer delivery delays. Note, however, that if large variations in incoming orders caused long delivery delays for the hoarders, the same large variations in incoming orders would cause long delivery delays for the phantom orders, because both $H$ and $P$ are involved in positive feedbacks (Fig. 2). Instead, there is no statistically significant difference in the variance of incoming orders for the PO group vs. the non-PO group. Nevertheless, follow up studies should seek other measures to provide additional tests of $H_0$ and $H_0$. 

---

9 Both $H_0$ (ii) and $H_0$ (ii) predict that the incidence of phantom ordering will not be associated with variability in incoming orders, though for different reasons. In fact, the standard deviation of incoming orders experienced by the phantom orderers is not statistically significantly different from that experienced by non-phantom orderers (two-tailed Mann-Whitney test, $U = 2697, z = 0.484, p = 0.63$). Variability in supplier reliability is associated with phantom ordering, but unexpected increases in incoming orders per se are not.

10 Because there are feedbacks among $S, SL^*$, orders and delivery delay (Fig. 2), it is possible that the association of hoarding with large peak delivery delays arises not because high delivery delays tend to trigger hoarding, but because the large orders

Please cite this article in press as: Sterman, J.D., Dogan, G., “I’m not hoarding, I'm just stockpiling before the hoarders get here.”. Behavioral causes of phantom ordering in supply chains. J. Operations Manage. (2015), http://dx.doi.org/10.1016/j.jom.2015.07.002.
resulted from heuristics such as “build safety stock to compensate for unpredictable customer demand” and “if suppliers deliver less than I need, order more to get what I really want.”

The results are striking because hoarding and phantom ordering are never rational in the experiment. There is common knowledge of the constant final customer demand. Participants also have common knowledge that there are no capacity constraints or random events, that cancellations are not permitted and that there is no horizontal competition—each supplier has only one customer and each customer has only one supplier. Participants who seek safety stocks or place phantom orders necessarily induce backlogs upstream and then accumulate excess inventory, raising costs for their supply chain and reducing their financial reward.

Why, then, do these individuals hoard and place phantom orders? Hoarding and phantom ordering are adaptive responses in many situations. Competition for limited supplies is common in settings where hoarding and phantom orders arise, such as shortages of “must have” toys and electronics, boom periods when suppliers are unable to fill orders on time or crises in which supplies of vital goods are inadequate and uncertain. Those who exercise restraint will often go without while more aggressive hoarders deplete available supplies. We speculate that participants exhibiting these behaviors in the experiment are inappropriately transferring heuristics for hoarding and phantom ordering from the real world, where they are sometimes adaptive, to the experiment, where they are dysfunctional.

Decision-making is not only a cognitive process, but is significantly affected by emotions. Environmental stressors often trigger autonomic arousal, including increased heart rate; adrenaline release, vasoconstriction that raises blood pressure, and pupil dilation, along with anxiety, fear, anger and other emotional and physiological reactions. A certain degree of autonomic arousal in response to stressors is adaptive and enhances survival (e.g., the “fight or flight” response to threat). But it has long been known that the relationship between arousal and performance is an S-shaped: too much stress worsens performance (Yerkes and Dodson, 1908). Causality is often two-way: stress may cause performance to suffer, and poor performance can increase stress (see Rudolph and Repenning, 2002). Thus, contrary to traditional economic theories where greater monetary rewards should induce greater effort and improve performance, higher stakes may increase stress to the point that performance suffers. Consistent with this hypothesis, Camerer and Hogarth (1999), reviewing the impact of incentives on performance in judgment and decision making tasks, find that incentives sometimes improve performance, sometimes have no impact, and sometimes decrease performance (they identify several reasons for this response; stress is only one). In the behavioral operations tradition, Bendoly (2011, 2013) found that participants in experimental real-time resource allocation tasks experienced significant arousal and stress, as measured by pupil dilation and blink rate, that these increased with task difficulty and that stress affected performance (see also Bendoly and Hur, 2007). Outside the laboratory, Lo and Repin (2002) measured the psychophysiological responses of professional securities traders during live trading sessions. They show that even the most experienced traders exhibit significant emotional and physiological responses to market events. Lo et al. (2005) also find that traders experiencing more intense emotional reactions to monetary gains and losses generated significantly lower financial returns.

In the CDKS implementation of the beer game, common knowledge that customer orders are constant means participants should expect no or only small, transient variations in incoming orders, and stable, timely deliveries from their suppliers. Deviations from these expectations are unexpected and costly surprises that may constitute stressors that increase autonomic arousal and heighten emotional responses to the point that they interfere with rational decision-making and trigger phantom ordering and hoarding.

The results support the situational hypothesis that stressors arising from scarcity or poor supplier delivery performance trigger hoarding and phantom ordering. The contrasting dispositional hypothesis posits that differences in the propensity to hoard or place phantom orders are due to idiosyncratic attributes of individuals, not the situation in which they find themselves. Although differences among individuals surely play a role, much of the response heterogeneity is explained by the conditions the participants faced, specifically the degree to which they experienced unexpectedly large orders or poor delivery performance.

People’s poor understanding of dynamics makes such irrational behavior more likely. Studies show that people have great difficulty accounting for feedback processes, time delays, and stock and flow structures that cause disequilibrium phenomena (Sterman, 2000 reviews the literature). Typical mental simulations of complex systems tend to assume open-loop, one-way causality, ignore time delays and do not account for accumulations (Cronin et al., 2009). Thus a firm may increase desired safety stocks and update their estimate of supplier lead times if suppliers miss a scheduled delivery. Such responses reflect open-loop, equilibrium mental models: having larger safety stocks and a larger supply line of unfilled orders is beneficial if customer orders are uncertain and suppliers are unable to deliver on time, but the only way to get them is to place more orders, further increasing the supplier’s lead time and the probability of missed deliveries. The stress caused by the deterioration in supplier performance may further degrade people’s ability to understand and respond appropriately to these disequilibrium dynamics, leading to still more phantom orders and eroding trust among supply chain partners in a destructive positive feedback.

The results, however, beg the question of why hoarding and phantom ordering are triggered for these individuals when they know it is irrational to do so in the experiment. Many animals, including humans, evolved in an environment characterized by scarcity, uncertainty, and competition for resources. During times of scarcity, such as winter, organisms must store resources sufficient to survive until food becomes abundant. Some species store the resources they collect as fat and hibernate during periods of scarcity. Others cannot use this strategy due to higher basal metabolic rates or the need to evade predators. For these species, including many birds and mammals, hoarding is highly adaptive (Grisham and Barlow, 2005; Batress and Day, 2003): individuals who do not “stock up before the hoarders get here” may not survive.

Brain imaging studies show that the neural structures responsible for the drive to collect objects, principally food, are located in the subcortex, a region of the brain that emerged early in mammalian evolution, while rational decision making in humans resides in the evolutionarily recent frontal cortex (Anderson et al., 2004). We speculate that the unexpected scarcity many participants experience in the experiment generates autonomic arousal sufficient to trigger the latent hoarding response, leading these individuals to boost their safety stock goals and place phantom orders. Doing so, of course, further degrades supplier service and may lead to still more stress and still more hoarding and phantom orders.

The psychiatry and neuroscience literatures provide support for these hypotheses. Hoarding is a common psychiatric disorder recognized in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA 2013). Symptoms include “excessive acquisition of possessions” and “persistent difficulty discarding or parting with possessions” due to “a perceived need to save the items and to distress associated with discarding them” (p. 247-248). Prevalence is estimated at 2-6% of the population. Hoarding disorder is related to Obsessive Compulsive Disorder (OCD) (APA 2013, Mataix-Cols and Pertusa, 2012; Bloch et al., 2008). Imaging studies demonstrate that lower activity in the dorsal anterior cingulate gyrus (ACC) is cor-

Please cite this article in press as: Sterman, J.D., Dogan, G., “I'm not hoarding, I'm just stocking up before the hoarders get here.”. Behavioral causes of phantom ordering in supply chains. J. Operations Manage. (2015). http://dx.doi.org/10.1016/j.jom.2015.07.002
related with hoarding among OCD patients. The ACG is associated with executive control, problem solving, selecting responses from multiple, conflicting options and emotional self-control (Saxena et al., 2004). Some individuals begin to exhibit compulsive hoarding after injury or surgery compromises the ACG or other regions of the frontal cortex associated with rational decision-making (Hahn et al., 2001; Eslinger and Damasio, 1985). Anderson et al. (2004) found that patients with no history of compulsive hoarding who suffered lesions in a particular region of the frontal cortex, specifically including the ACG, were far more likely to engage in compulsive collecting behavior afterward than those with lesions in other regions of the frontal cortex. They conclude that the injured region “disrupts a mechanism which normally modulates subcortically driven predispositions to acquire and collect, and adjusts these predispositions to environmental context.”

It might be argued that evidence about individuals suffering from hoarding disorder and patients with brain injuries does not apply to the healthy participants in the experiment or to managers of real supply chains. However, showing normal individuals photographs suggestive of hoarding, such as large stacks of old newspapers, provokes urges and activates brain centers similar to those observed in compulsive hoarders (Mataix-Cols et al., 2003; Phillips and Mataix-Cols, 2004), with the latter finding that presentation of hoarding-related images also led to “significant increases in anxiety” (see also An et al., 2009). Many participants in the CDKS experiment experienced cues suggestive of scarcity, including large, unexpected increases in incoming orders, large backlogs, and delayed deliveries from their supplier. We speculate that such cues activate the hoarding impulse in the subcortex, while at the same time acting as stressors enhancing autonomic arousal that may suppress the ability to regulate that impulse.

The results suggest a variety of avenues for further research. The hypothesis that those engaging in phantom ordering and hoarding experienced more autonomic arousal than others can be tested directly as people participate in the experiment, as Lo and Repin (2002) did for traders. The hypothesis that stressors such as large incoming orders and supplier shortages suppress the ACG and/or activate hoarding centers in the subcortex or limbic system can be tested by fMRI or PET imaging of participants during the experiment. For these purposes it may be desirable to use a simplified inventory management system in which incoming orders and supplier lead times can be manipulated directly, whereas in the full BDG these are endogenous.

Empirical studies of supply chains in which instability is prevalent may also be fruitful. Customers and managers often experience long and unpredictable supplier lead times, stockouts, and other cues that signal scarcity and may constitute stressors that could trigger hoarding or phantom ordering. Such experiences may be far more stressful than the experiment. Interviews and econometric studies may document the prevalence of such behavior. Interviews, however, are unlikely to reveal the extent to which hoarding and phantom ordering are strategic, rational responses or emotional reactions to scarcity, since informants may seek to portray themselves in a favorable light or suffer from hindsight bias and ex post rationalization. Observation of supply chain managers in real time during periods of scarcity may provide better tests of the role of emotions, panic, and fear. Similar experiments and field studies in related contexts including speculative bubbles, bank runs, riots and escape panics may also shed light on the neurological, physiological, emotional and strategic processes affecting behavior.

The results also suggest normative implications. To the extent everyone is susceptible to hoarding and phantom ordering to some degree, with these behaviors triggered by stresses in the situation, firms must not only attend to the game-theoretic structure of their supply chains but must also design information systems and other features of the workplace to minimize the stresses that might trigger dysfunctional behavioral responses to scarcity. Studies exploring how to reduce these stressors, whether and how training can help people resist the urge to hoard, and how to design automated supply chain management systems to reduce their vulnerability to hoarding and phantom ordering may help stabilize supply chains currently at risk in both business and public policy contexts.

Finally, the results suggest a greater role for behavioral research in operations management generally. Hoarding and phantom ordering are not the only behavioral dynamics relevant to manufacturing and service operations. Other settings where behavioral reactions are likely to play a role include the hockey stick in sales and manufacturing, cost and schedule overruns in project management, and firefighting in new product development, to name just a few (e.g., Lyneis and Ford, 2007; Repenning and Sterman, 2002: Repenning, 2001). Effective behavioral research in operations will unite the tools of traditional operations management and operations research with organization theory, psychology, neuroscience, system dynamics and other disciplines. The result will be rich behavioral theories of disequilibrium dynamics, tested in experiments, fieldwork and other empirical studies, and generalized with formal models that offer useful insights to improve the management of complex operations.

Acknowledgments

We are grateful to Rachel Croson, Karen Donohue, and Elena Katok for providing the data analyzed here, and for helpful comments; we also thank Elliot Bendoly, Kevin Linderman, Rogelio Oliva, the referees, and many others who provided useful suggestions in seminars and conference presentations. We thank the Project on Innovation in Markets and Organizations at the MIT Sloan School of Management for financial support.

References


abstract=2286492 or http://dx.doi.org/10.2139/psp Shapiro, D., Green, G.A., 1977. The anti-inflammatory properties of
sulfuric acid. In: Proceedings of the Conference on the Use of
Sulfuric Acid in Medicine and Agriculture. The American
Stern, J.D., Dogan, G., 2010. Behavioral causes of phantom ordering in
supply chains. J. Operations Manage. 28 (3), 129–156.