The bullwhip effect

Fransoo, Jan C.; Udenio, Maximiliano

Published in:
International Encyclopedia of Transportation

DOI:
10.1016/B978-0-08-102671-7.10232-5

Publication date:
2021

Document Version
Peer reviewed version

Citation for published version (APA):
Fransoo, J. C., & Udenio, M. (2021). The bullwhip effect. In R. Vickerman (Ed.), *International Encyclopedia of Transportation* (Vol. 3, pp. 130-135). Elsevier. https://doi.org/10.1016/B978-0-08-102671-7.10232-5

Bullwhip Effect

Maximiliano Udenio & Jan C. Fransoo

May 2019

1 Definition

The bullwhip effect is a supply chain phenomenon comprised of two information-distortion mechanisms. The first is that, for a given firm, the orders it places to its suppliers tend to be more variable than the demand it observes from its customers. This is called *demand distortion*. The second mechanism is the observation that demand distortion increases the further upstream a firm is in its supply chain (i.e., further away from the final consumer). This is called *variance amplification*.

Due to the combined effects of demand distortion and variance amplification, a demand shock downstream generates demands that oscillate with increasing amplitude at each successive stage of the supply chain. This is said to resemble the visual effect of a bullwhip, with which a cattle farmer can use a flick of the wrist to break the sound barrier; hence the name. Figure 1 illustrates this resemblance, showing the evolution of demand on a stylized supply chain following a sudden, 2 percent, change in the end market.

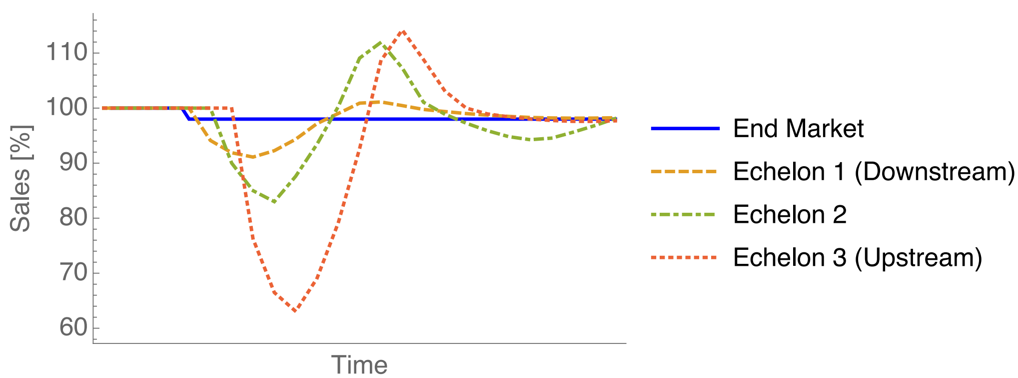

Figure 1: Illustration of variance amplification of demand on successive supply chain echelons

Even though the term itself dates from the late 90's (see Lee et al., 1997a,b, for the earliest appearances of the term), the mechanisms behind the bullwhip effect were first described in the late 50's by Jay Forrester, a professor at Massachusetts Institute of Technology. The bullwhip effect is, thus, also frequently referred to as the "Forrester effect", named after his seminal initial work (Forrester, 1961). The bullwhip effect is often described in research through stylized models, typically using a non-stationary downstream demand stream (e.g., a sudden shock) as a trigger for instability. The bullwhip effect, however, can also be observed under stationary demand.

One of the most significant contributions of the work of Lee et al. (1997a) was to show, mathematically, that the bullwhip effect can appear even under conditions of full rationality. Thus, while irrational behavior can increase its impact (e.g., Sterman, 1989; Croson and Donohue, 2006), its appearance is not limited to improperly managed firms.

The bullwhip effect is undesirable because it generally increases operational costs and may cause further management problems in the supply chain, such as poor service levels. Demand distortion leads to inefficiencies in production and inventory management: Highly variable production and/or inventories are required to serve highly variable (oscillatory) customer demands. Therefore, the bullwhip effect is directly responsible for costs associated with starting and stopping production equipment, successive idling and overtime periods, and/or workforce changes. Moreover, the bullwhip effect can also have a negative impact in ways that are difficult to quantify, such as difficulties in forecasting demand, and customer relations problems such as stockouts or delays in supply (Disney et al., 2013).

Because of variance amplification, the bullwhip effect is specially problematic for firms that are upstream in the supply chain. The compounding effect of variance amplification increases the impact of the bullwhip effect on firms located upstream a supply chain; a chemical company, for instance, will typically suffer a larger bullwhip than a retailer. In fact, the inefficiencies associated by the progressive increase in upstream variability have a measurable effect on the relative production and inventory costs for upstream firms. Some studies estimate that these inefficiencies account for excess costs between 10% to 25% (Lee et al., 1997a).

From an information perspective, the importance of the bullwhip effect lies in that it makes it increasingly difficult for firms to extract meaningful information from demand data. As one travels upstream a supply chain, demand signals progressively diverge from the original consumer end-market demand. It is possible that after a few stages, very little resemblance remains (Udenio et al., 2015). Therefore, it is difficult for upstream firms to use their own demand signals to estimate consumer demand and to account for changes in the consumer market. In fact, most common examples and case studies based on the bullwhip effect illustrate this particular aspect; e.g., how the production of diapers is more variable than what would be expected from the (relatively stable) usage by babies and toddlers, or how the the DRAM market faces a much higher volatility than the computer market.

This is of particular importance when consumer demand experiences a shock and when no information is shared among firms. In fact, minor consumer demand shocks result in upstream production levels that oscillate around consumer demand levels; making it impossible for upstream firms to link their demand observations with downstream market behavior merely based on the demand signal. This lack of visibility makes accounting for the bullwhip effect (i.e., taming the bullwhip) a difficult problem. Depending on the structure of the supply chain, downstream shocks can propagate and the related oscillations can persist for years.

In the academic literature, the bullwhip effect is classified in numerous ways. In this entry, we analyze the bullwhip effect through three different theoretical causes. Namely, the bullwhip effect caused by *delays* within the supply chain structure (e.g., Forrester, 1961), the bullwhip effect caused by *behavioral biases* (e.g., Sterman, 1989), and the bullwhip effect caused by rational *decision-making* policies (e.g., Lee et al., 1997a). Whereas any one of the above is able to explain the appearance of the bullwhip effect, it is typical – in real life – to find more than one working concurrently.

The bullwhip effect is commonly introduced to students and executives in the classroom using the *beer distribution game* developed by John Sterman. In the beer distribution game (see Sterman, 1989, for a detailed overview) four players (or eventually, groups of players) need to coordinate a supply chain consisting of four echelons, without any interaction among the echelons being allowed. Commonly, this leads to excessive variance amplification over the course of the game. The bullwhip effect observed during a typical session of the beer game is caused by a combination of all the causes we discuss in this entry: the delay structure of the supply chain makes the oscillations all but inevitable, whereas the behavioral biases of the players, as well as a number of the structural causes, exacerbate the effect.

It is important to note that the bullwhip effect is at odds with the economic theory of production smoothing, whereupon production is kept constant in time and demand fluctuations are absorbed entirely through inventory (Abramovitz, 1950). The theoretical justification behind production smoothing is the assumption that changes in production output are generally more expensive than the necessary inventory buffers, thus it would be rational for firms to maintain production as stable as possible; allowing inventory levels to increase when sales decrease and vice versa. Successive steps of such countercyclical behavior implies that in a supply chain the variability of upstream demand would be lower than the variability of the end-market.

Production smoothing was a common assumption in macroeconomic modeling until the

80's. This modeling paradigm, however, is difficult to justify empirically (Blinder and Maccini, 1991). At a macro-economic level, inventory investment is strongly procyclical and, as discussed, production tends to be more variable than sales. To accommodate these observations, a number of economists have expanded the production smoothing model, noting that the addition of cost and/or productivity shocks can reconcile the model to observations. A different view in macroeconomic modeling eschews the production model altogether, arguing that avoiding stockouts is the more important metric for a firm, and that procyclical behavior and a production variability that is larger than demand variability (i.e., the bullwhip effect) follow naturally from this assumption (Kahn, 1992).

2 Causes of the Bullwhip Effect

The bullwhip effect was first analyzed in the late 50's as a product of inventory and ordering policies and inherent delays in the supply chain structure. In the late 80's, classroom experiments using the beer distribution game were conducted and human biases were introduced to extend the understanding and analysis. In the late 90's, a more mathematically rigorous analysis of the bullwhip effect showed that supply chain decision-making, without the impact of delays or human biases is enough to trigger the bullwhip. This triggered extensive research work encompassing more formal modeling, more empirical work at firm, supply chain, and industry sector levels, as well as more laboratory experiments detailing out earlier understanding.

2.1 Delay structure

In the late 50's, Jay Forrester introduced "industrial dynamics" (Forrester, 1961), and with it a new modeling paradigm for production and inventory systems. In this view, the defining factor causing the oscillatory dynamics typical of the bullwhip effect is the interaction of inventory policies and delays.

Delays induce the bullwhip effect because they disconnect decisions from the observation of their results. When there is a delay between making a decision and the observation of its result, then every subsequent decision must consider this delay and the associated uncertainty. Operationally, this implies keeping track of decisions taken but whose results are not yet observed and basing these decisions on expectations of future realizations. A typical example of a delay in a supply chain is the replenishment lead time; if the replenishment lead time is longer than the ordering frequency, then this delay forces the decision maker to keep track of multiple orders and to predict demand multiple periods ahead. Such delays complicate the analysis, noting that not all information in the supply chain can be observed and future demand is unknown.

Assuming there is a desired *end state* for a system, all decisions must be made in function of this result. In practice, this implies that decision makers must decide, not only on desired inventory and production levels, but also on the path towards them. This is modeled through adjustment times; the desired amount of time that a decision maker wishes to spend chasing the desired state. Implicitly, this assumes a policy-maker, and policy-related decisions. Such policies, the underlying behavior, and their effect on the bullwhip effect are explicitly modeled in the behavioral operations literature.

2.2 Behavioral causes of the bullwhip

Behavioral models of the bullwhip effect are commonly based upon the dynamic models introduced in Forrester (1961). Behavior is explicitly modeled through adjustment times—short adjustment times implying 'nervous' behavior, prioritizing quickly reaching the end state at the expense of variability in orders/production, and long adjustment times as a way to model 'relaxed' behavior, prioritizing stability over time to reach the desired state.

Behavioral models of the bullwhip effect are commonly investigated in experimental research using the beer distribution game (Sterman, 1989; Croson and Donohue, 2006). In this game four players assume different roles in a supply chain, from retailer to beer producer. In its basic form, no communication is allowed between players and an unknown consumer demand must be filled by the retailer, who can place orders to its supplier, who can place orders to its supplier, and so on, until the beer producer places manufacturing orders. The beer distribution game introduces information and material delays into the system, i.e., it takes time for order information to travel upstream, and it takes time for material ("beer") to travel downstream. Analysis of data from the beer game shows that, in addition to the delays, behavioral attributes increase the magnitude of the bullwhip effect.

The main behavioral insight to come out from the beer game is that humans are not good at accounting for the pipeline (i.e., everything that has been ordered but not yet received). Compounded with long delays, this causes players to overreact to stockouts and order every period without taking into account outstanding orders. The effect of ignoring the pipeline is often seen as enormous oscillations, deep stockouts followed by equally large excess inventory. This effect is very robust and has been replicated innumerable times (e.g., Croson and Donohue, 2003; Croson et al., 2014; Narayanan and Moritz, 2015; Moritz et al., 2019)

Being based on the same paradigm as Forrester models, the underlying cause of the bullwhip effect in the beer game is the delay structure of the system. The insight from experimental research lies in identifying the behavioral biases present and how this irrational behavior further amplifies the bullwhip effect.

2.3 Rational decision-making causes

Since the publication of the work by Lee et al. (1997a,b), there has been significant interest in the theoretical mechanisms causing the bullwhip effect under fully rational decision makers. Under a periodic review inventory policy, for a firm calculating the optimal order quantity every period, the bullwhip effect appears due to four fundamental causes: Demand signal processing, order batching, price fluctuations, and shortage games. These insights have been obtained using stylized analytical models.

Demand signal processing. When a firm uses past demand information to forecast future demand, a non-stationary change in demand (e.g., a one-period surge or drop in demand) will trigger the bullwhip effect.

In the absence of communication among firms, a supplier uses observed demand to update its forecasts. In such situations, non-stationary customer demands trigger changes in supplier forecasts that lead to increased variability due to a constant update of the optimal order quantity. The amplification of demand variability increases with the lead time, but it has been shown in an analytical model that amplification exists in this setting even when lead times are zero (Lee et al., 1997a).

Order batching. If there is a non-zero ordering cost associated to placing an order, then order batching occurs naturally. If a supplier has a single customer order batching implies that it will observe less frequent, larger orders, i.e., increased variability. If a supplier has multiple customers, the impact of order batching depends on the correlation and timing of the individual demand streams. Under "perfectly synchronized" retailer ordering, i.e., when a constant number of customers order every period, then the contribution of order batching to the bullwhip effect is avoided. In all other cases, order batching generates a bullwhip effect. Positively correlated ordering results in higher amplification than uncorrelated (i.e., random) ordering, which in turn results in higher amplification than balanced demands. Balanced demands are those whereupon groups of customers place orders within a designated period every review cycle without any overlap (unlike perfectly balanced demands, the number of retailers placing orders in a given period is not constant).

Price fluctuations. When a firm experiences variations in the purchase price, then it is optimal to adopt an ordering policy such that there exist different order-up-to levels, decreasing in the purchase price. All else equal, this results in increased variance of orders and hence a bullwhip effect, compared to a policy with constant prices. As an example, consider a product with two prices, the regular price and a sale price. Unless the buying firm has full advance information on the sales periods, it follows logically that the optimal order-up-to level will increase whenever there is a sale and decrease when the price returns to normal.

Shortage gaming. Shortage gaming appears when a firm orders from a supplier that regularly suffers from stockouts. Facing a shortage, a supplier must allocate the available inventory among its customers. If such an allocation is (or is thought to be) proportional to the order size, customers have an incentive to inflate orders in an effort to obtain a larger share of the allocation. This sends a false demand signal that, interpreted by the supplier as a real increase in the demand, compounds the amplification problem (see demand signal processing above). In addition, customers that inflated orders will typically cancel superfluous orders once the supplier catches up with demand; leading to further increase in demand variability.

3 Measuring the bullwhip effect

3.1 The bullwhip in real life – anecdotal evidence

Anecdotal evidence of the bullwhip effect abounds. In addition to the examples mentioned above, other famous cases are often used as case studies in business schools and the practitioner literature. Canned soup, dried pasta, and printer ink are some of the well-known examples of evidence of the bullwhip effect—products with stable consumer demand and highly variable upstream production. In recent decades, the bullwhip effect has become shorthand for the 'inevitable' variability that upstream firms are required to contend with.

Nevertheless, for all the usefulness of anecdotal evidence and case studies, rigorous research requires rigorous evidence of the existence of the bullwhip effect as a widespread phenomenon.

3.2 The bullwhip in real life – empirical evidence

For all our understanding of the theoretical and behavioral causes of the bullwhip effect, extensive empirical research has been and is still conducted to show whether the bullwhip effect is an industry-wide problem, or it only affects a limited number of individual firms. To this end, numerous researchers have attempted to rigorously measure it using large panels of data. The evidence for the bullwhip effect in empirical data is not clear cut (see Cachon et al., 2007; Bray and Mendelson, 2012, for discussions of why this is the case). There are a number of reasons that explain the nuances observed in the research.

First, empirical research–whether it uses primary or secondary data–typically uses aggregated data. It has been shown mathematically that measuring the bullwhip effect in aggregate data underestimates its impact (Chen and Lee, 2012). The particular degree of aggregation varies from study to study. Most empirical bullwhip effect studies use secondary data aggregated at the firm level with regards to products (i.e., all products are aggregated into a single stream), to inter-firm links (i.e., all customers/suppliers are aggregated into a single stream), as well as at the temporal level (observations are aggregated by month, quarter, or even year) (Fransoo and Wouters, 2000).

Second, empirical research uses proxies to measure certain quantities. For example, sales data is used as a proxy of demand data, and is combined with inventory data to estimate production data. In addition to the inherent errors that such estimations can introduce, the use of such data is conceptually at odds with the theory of the bullwhip effect. This is because the theoretical development of the bullwhip effect relies on the use of information flows in addition to material flows. As an example, the implicit assumption behind the modeling of the delay structure of a supply chain used in the beer game is that lead times are the result of a combination of the delay in information transmission (e.g., the time between placing an order, and a supplier executing the necessary steps) and the material delay (e.g., the time it takes an order to be physically transported from a firm to its customer). Chen et al. (2016) show that using material flow data to estimate the bullwhip effect can over- or underestimate the *real* bullwhip effect; that is, the bullwhip effect as measured using information flow.

Another nuance that appears when measuring the bullwhip effect on empirical data is the stabilizing effect of demand seasonality. In real life, a significant number of products exhibit demand seasonality to a certain extent. For researchers estimating the bullwhip effect observed in the real data, the question is, then, whether this should be measured using the original demand stream, or a deseasonalized demand stream (Cachon et al., 2007; Chen and Lee, 2012). From a theoretical perspective, including seasonality in the measurements can dampen the observed bullwhip; particularly if the variability of seasonality dominates the variability of the demand itself, and more so if there are capacity constraints.

Of course, another source of ambiguity in the measured data is the fact that many firms actively seek to limit the bullwhip effect, thus, in reality there is a mix of bullwhip effect and production smoothing and disentangling the one from the other is non-trivial.

3.3 Evidence from real macro-economic shocks

A number of studies took advantage of the 2008 credit crisis as a quasi-natural experiment, where the demand collapse in multiple industries approximated a non-stationary shock across entire industries. Evidence of an inventory shock is particularly strong. Such studies support the predictions regarding the transmission of the variability. In general upstream firms have observed more severe drops in demand than downstream firms. Moreover, recent studies claim that the inventory shocks by themselves – i.e., independent as a response to changing demand, but as a consequence of general financial or economic conditions – further amplify the bullwhip effect (Udenio et al., 2015).

4 Taming the bullwhip

It is impossible to eliminate all structural and behavioral causes of the bullwhip from real life supply chains. Thus, decision-makers can attempt to, at best, "tame" the bullwhip as much as possible. Ample research exists prescribing strategies to "tame" the bullwhip. Each of these strategies tackles one (or more) of the causes of the bullwhip effect highlighted above.

Delays in the supply chain structure amplify variability; thus, minimizing delays is an obvious strategy to reduce amplification. Clearly, firms can only reduce delays by altering the structure of the supply chain in which they operate. All else equal, a supply chain with short lead times will experience a smaller bullwhip effect than a aupply chain with very long lead times. Firms willing to limit the appearance of the bullwhip are recommended, to the extent possible, to participate in agile supply chains; source (sell) from (to) suppliers (customers) with short lead-times. In practice, local supply chains are typically more agile than global supply chains. Note that the mode of transport used for deliveries affects lead times and thus the bullwhip effect. Thus, such a strategic decision must consider the impact on variability amplification in addition to trade-off between logistics costs, lead times, and environmental impact.

From a behavioral standpoint, the main recommendation is to avoid under-weighing the pipeline. While this recommendation directly addresses the main behavioral bias exhibited by human players of the beer game, research has shown that firms operating with policies that track "desired" inventory and pipeline targets (e.g., order-up-to policies) are also prone to under-weighing the pipeline (be it by design or through arbitrary adjustments). Further, adopting fractional–rather than full–adjustments decreases the bullwhip effect generated by such policies. Such fractional-policies, however, reduce the bullwhip at the expense of responsiveness to changes in demand (Disney et al., 2013).

Information sharing is often prescribed as a tool to counter the bullwhip effect because it tackles a number of structural issues. Under full information sharing, all firms in a supply chain can observe up-to-date inventory and/or demand information for all other firms in the supply chain. The visibility afforded by information sharing breaks the main mechanism behind the demand forecast updating; in this setting firms can generate more reliable forecasts, for example by using downstream demand and inventory information to estimate future demand. This has been shown to be a particularly powerful action when combined with knowledge about the inventory policies that other firms employ. Directly sharing demand (and future order) forecasts among firms can similarly reduce the bullwhip effect (**?**).

In addition, information sharing can also mitigate the effect of shortage gaming, on the one hand by allowing suppliers to understand the "real" demand from customers (and thus avoid reacting to artificially inflated orders) and on the other hand by providing customers with information regarding the shortages that a supplier is facing. If a customer knows about the severity of a particular shortage, as well as the allocation mechanism, and the expected recovery time, the incentive to artificially inflate orders to increase the allocation is greatly reduced. Full information sharing across a supply chain was previously considered all but unimplementable due to operational and technological barriers. Today, information sharing is starting to become technologically feasible. However, a strong barrier its adoption is that, oftentimes, firms see demand and inventory information as a source of competitive advantage and thus classify it as sensitive, and are not willing to share. Furthermore, the immediate benefits of information sharing are not symmetric. In fact, taming the bullwhip through information sharing requires downstream firms (the least affected by the bullwhip) to share information so that upstream firms reap the benefits. Downstream firms do not get comparable benefits from obtaining upstream demand/inventory information. Thus, many downstream firms do not see this as a fair trade-off. This reasoning, however, negates the second order benefits that downstream firms are able to obtain; if upstream companies achieve bullwhip reduction, then the decrease in inefficiencies and associated costs will sooner or later reach downstream firms in the form of lower costs, increased reliability and shorter lead times. Hence, information sharing typically needs to be associated with incentive alignment to be effective.

Other operational policies are designed to tame the bullwhip by attacking the root of each of the structural causes. For example, adopting "every day low prices" policies, or pricing contracts, to avoid the variability generated by price fluctuations; encouraging less-than-full truckload ordering via discounts and arrangements with 3PL's to discourage order batching; and adopting allocation rules that are independent of current orders (e.g., allocating based on past sales) to prevent rationing games.

Finally, there is the case for vertical integration. Information sharing is widely viewed as a way forward in regards to synchronizing the management of multiple firms, and thus reducing the bullwhip effect. As discussed above, such implementations are relatively rare to find at large scale due to the view of information as competitive advantage. Vertical integration of firms enables what in practice is full information sharing of different stages of a supply chain while at the same time avoiding competitive issues.

References

Abramovitz, M. (1950). Inventories and business cycles, with special reference to manufacturer's inventories. National Bureau of Economic Research.

Blinder, A. and Maccini, L. (1991). Taking stock: A critical assessment of recent research on inventories. The journal of economic perspectives, 5(1):73–96.

Bray, R. L. and Mendelson, H. (2012). Information transmission and the bullwhip effect: An empirical investigation. Management Science, 58(5):860–875.

Cachon, P., Randall, T., and Schmidt, G. (2007). In search of the bullwhip effect. Manufacturing & Service Operations Management, 9(4):457.

Chen, L. and Lee, H. L. (2012). Bullwhip effect measurement and its implications. Operations Research, 60(4):771–784.

Chen, L., Luo, W., and Shang, K. (2016). Measuring the bullwhip effect: Discrepancy and alignment between information and material flows. Manufacturing & Service Operations Management, 19(1):36–51.

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

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

Disney, S. M., Hoshiko, L., Polley, L., and Weigel, C. (2013). Removing bullwhip from lexmark's toner operations. In Production and Operations Management Society Annual Conference, pages 3–6.

Forrester, J. W. (1961). Industrial dynamics. MIT Press (Cambridge, Mass.).

Fransoo, J. and Wouters, M. (2000). Measuring the bullwhip effect in the supply chain. Supply Chain Management: An international journal, 5(2):78–89.

Kahn, J. A. (1992). Why is production more volatile than sales? theory and evidence on the stockout-avoidance motive for inventory-holding. The Quarterly Journal of Economics, 107(2):481–510.

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

Lee, H. L., Padmanabhan, V., and Whang, S. (1997b). The bullwhip effect in supply chains. Sloan management review, 38:93–102.

Moritz, B., Narayanan, A., and Parker, C. (2019). Unraveling behavioral ordering: Relative costs and the bullwhip effect. Available at SSRN 3195282.

Narayanan, A. and Moritz, B. B. (2015). Decision making and cognition in multi-echelon supply chains: an experimental study. Production and Operations Management, 24(8):1216–1234.

Sterman, J. (1989). Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. Management Science, 35(3):321–339.

Udenio, M., Fransoo, J., and Peels, R. (2015). Destocking, the bullwhip effect, and the credit crisis: empirical modeling of supply chain dynamics. International Journal of Production Economics, 160:34–46.