Inventory dynamics and the bullwhip effect : studies in supply chain performance

DOI:
10.6100/IR776508

Document status and date:
Published: 01/01/2014

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Inventory Dynamics and the Bullwhip Effect
Studies in Supply Chain Performance

Maximiliano Udenio

Printed by Proefschriftmaken.nl || Uitgeverij BOXPress

Cover design by Victor Anelli

This thesis is number D184 of the thesis series of the Beta Research School for Operations Management and Logistics. The Beta Research School is a joint effort of the School of Industrial Engineering and the Department of Mathematics and Computer Science at Eindhoven University of Technology, and the Center for Production, Logistics and Operations Management at the University of Twente.

A catalogue record is available from the Eindhoven University of Technology Library.

ISBN: 978-90-386-3664-1

This research has been partially funded by the Dutch Institute for Advanced Logistics (DINALOG), within the context of the 4C4Chem project: Cross-Chain Collaboration in the Chemical Industry.

Inventory Dynamics and the Bullwhip Effect
Studies in Supply Chain Performance

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. C.J. van Duijn, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op donderdag 25 september 2014 om 16:00 uur

door

Maximiliano Udenio Castro

geboren te Buenos Aires, Argentinië

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr. J.W.M. Bertrand
1e promotor:	prof.dr.ir. J.C. Fransoo
2e promotor:	prof.dr. A.G. de Kok
leden:	dr.ir. N.P. Dellaert
	prof.dr. V. Gaur (Cornell University)
	prof.dr. K. Hoberg (Kühne Logistics University)
	prof.dr.ir P.M.J. Van den Hof

"I love deadlines. I like the swooshing sound they make as they fly by."

Douglas Adams

Acknowledgments

I had a lot of deadlines these past four years. Some of them were self imposed, some of them were not. Some were super-important, some not so much. But through it all, always in my head was this *big one. The* deadline. The *finish your dissertation* deadline.

I stressed about this deadline quite a bit, especially during the first year of my Ph.D. *How the hell am I going to produce enough original research?* Worse, *how am I going to write all that down?* Well, at some point I just stopped worrying about the big picture, and concentrated in going step-by-step; deadline-by-deadline. In hindsight, it's a great strategy for completing a long project like this, and I would love to say it was all part of a plan. However, I did it mainly because I could not afford to worry about the dissertation on top of everything else—it would've driven me mad. The dissertation itself never, ever, left the back of my mind, but I sort of learned to live with that feeling—some sort of survival instinct, I guess. But It works, and by, say, the middle of year 4 I 'suddenly' realized that I had, somehow, accumulated quite a lot of work. Not only that, but when taken together, it started to look like a dissertation. There was a *thesis* behind it, and the only thing left for me to do was to discover it and shape it. And stress out over layout and fonts and whatnot.

My point is that this dissertation is the result of 4 years of constant tinkering and ruminating on ideas; 4 years of long, short, planned, improvised, fulfilling, grueling discussions with lots of people; 4 years of writing down these ideas little by little—of course, the best ideas usually came in comically inconvenient times: In the shower, in dreams, biking in the middle of nowhere... everywhere, it seems, except while at my desk. This dissertation is not the result of me going into a deep cave with some blank paper, waiting for a flash of inspiration to take over me, only to emerge several months later with a bunch of finished books. What I'm trying to say that there's no way I could've done this on my own. I owe profound gratitude to so many people. You've all been so patient and generous and understanding. Let me start with Jan.

None of this would have happened without prof.dr.ir. Jan Fransoo. Not only because he is the one who actually hired me for this position, but because without his guidance I wouldn't have made it past the first tiny step. Jan pushed me when I needed a push, held me back when I was about to go down some deep rabbit-hole, and always –always– asked the right questions. The kind of questions that help you turn confusion into research. He played a crucial role in every chapter of this dissertation. I lost count of the number times where I went to Jan with some 'enormous' problem, thinking all is lost, only to see him chuckle, think for 5 seconds and solve everything with some short and astute observation. On the flip side, but just as important, were all those times where I proudly went to Jan with some 'genius' idea, thinking I'm so clever, only to see him chuckle, think for 5 seconds and disintegrate my silly idea with some short and astute observation. Thank you Jan, for pushing me and for holding me back.

Prof.dr. Ton de Kok is my second promotor. It's always fun being around Ton. He shares his views on our field, and research in general, with such passion that it is invigorating. He made me think about what it all means, he helped me grasp the big picture; made me take a step back and look at everything I've learned as different parts of a whole. Thank you Ton for your ideas and suggestions.

Professor Vishal Gaur, from the Johnson School of Business at Cornell University, hosted me for three months in beautiful Ithaca, New York. I worked together with Vishal in what is now Chapter 6. I cannot express how grateful I am for this opportunity; those three months were an insanely great experience. Vishal gave me so much of his time that I seriously couldn't believe it. He guided me brilliantly, encouraging me all of the time. I hope that this collaboration is but a start. I also want to thank Vishal for accepting being part of my doctoral committee and helping improve my dissertation with his comments.

The work contained in Chapter 4 is a collaboration with Eleni Vatamidou and dr. Nico Dellaert. Without Eleni, Theorem 4.1 would be conjecture 4.1. We worked long nights and weekends together, chasing after that elusive proof. Thanks Eleni for being awesome. For all the food and coffee; for your crusade against long, confusing sentences that only I understand. Nico helped to make this chapter be much better. His comments after reading that somewhat confusing first draft helped me find the focus it needed. I would also like to thank him for agreeing to be a part of my doctoral committee and for his comments on the rest of the work.

I first met prof.dr. Kai Hoberg (from the Kühne Logistics University, Hamburg) not in person but by dissecting and analyzing his control theory papers—his influence in Chapter 4 is huge. We would later meet in real life and have the opportunity to work together. I want to thank Kai for being part of my committee. He brought his experience in both the fields of control theory and empirical research, which resulted in a stronger dissertation all around.

I would also like to thank the final member of my doctoral committee, prof.dr.ir Paul van den Hof from the department of electrical engineering at the TU/e. His profound knowledge of control systems made me think back reflectively and make several changes that strengthened chapters 3, 4, and 5.

Chapter 3 is product of a collaboration with Robert Peels, from Flostoc BV. I want to thank Robert for the many, many interesting conversations we've had over the years. His enthusiasm, and the drive he has to go from very abstract ideas to something practical are contagious.

I also want to thank each and every person I had the chance to meet at OPAC over the years. Thanks for all the fun chats, the (sometimes absurd, always interesting) discussions, and all the conferences we've shared. Also for all the birthday cakes and cookies. To all the secretaries, thanks for all the hard work we take for granted. To the 'old guard' of OPAC Ph.D's, who were about to finish when I started, thanks for the advice. Especially Ola, Kurt, Said, and Ben. I would also like to thank all of those with whom I shared an office through the years: Duygu, Derya, Baoxiang, Frank, Kristel, and Hande. Anna, Yann, José, Josue; thanks for the fun times. Qiushi, Joachim, Chiel, Maryam, Stefano, Maarten, Frank; thanks for the tea-break, and wine hour, discussions. Zumbul and Tarkan, thanks for all that coffee and advice. Kasper: man, it's actually doable, see? Your turn now. There are so many people with whom I shared corridors E and F, I cannot possibly name you all—thank you.

To my friends. Those back home, Papu, Padi, Quelo, Coc, Boti, Bolo, Guille, Ale, Diegui, Mago, Conexo, Juanma, Juanito, Tommy, Rudo (you blew my mind with your cover design). To those abroad, Chicky, Marto, Lucho, Frodo, Santi, Guille. I don't see enough of you guys. You are always in my thoughts. Gauchito, I miss you dearly.

To my "dutch" friends. Brian, my brother away from home. Thank you for being there for me, you never stopped being the "adult in charge" for me. Vseva, thanks for all the bikes and all the chess. Giorgos, Maria, Dima, thanks for the trips, and all the food. To all of you: I'm sorry for my anti-social phase. I hope you understand what I was doing all those weekends/evenings that I would just ignore your messages and calls.

To my family? Well, I would literally not be here today without them. Mamá y papá, los extraño y quiero tantísimo. Gracias por entender porqué me vine a Holanda. Se cuán difícil es para ustedes. No me alcanzaría toda la tinta del mundo para agradecerles. Gracias por estar siempre. Gracias tías y tíos, Fanny, Alfre, Ali, Silvia, Lina, por hacerme sentir tan querido cada vez que estoy por allá. Bruno, my big brother. As much as I hate to admit it (and won't probably do in public again) you did and do so much for me. I can't imagine living without you covering for me. Thank you. My beautiful nieces: Catalina, Emilia y Sofia. I hope some day you pick up this little book and get to understand why I wasn't there while you were growing up. I miss you more than you know. Nati, not only are you the mother of those three cuties, but you are the big sister I never had. I know I puzzle you sometimes, and that my uncommunicativeness drives you up the wall, and yet, you are always there. Thank you.

One last word for Heather. Your love and support carried me through this all. I truly couldn't have done it without you. Tons of love. Thank you Heather. Theather.

Maximiliano Udenio
Eindhoven, August 2014.

Contents

> If the pages of this book contain some successful verse, the reader must excuse me the discourtesy of having usurped it first. Our nothingness differs little; it is a trivial and chance circumstance that you should be the reader of these exercises and I their author.
>
> Jorge Luis Borges

Chapter 1 Introduction

Look around you. Virtually everything you see was at some point stored as 'inventory', awaiting for its transformation or sale. Everything you see was produced, and stored, according to a policy that a human put in place.

Not too long ago, I contacted a company in the Netherlands to print my dissertation. The order I placed affected their paper and ink and glue stocks; these raw materials were transformed into the convenient book you are holding right now. The decision of how many books to get printed was not trivial: Had I made a rushed decision and ordered too few of them, and you may have not had the chance to be reading this right now; had I ordered too many of them, and I would have faced a problem: *where do I put all these boxes and boxes of excess dissertations?* In the end, I ordered enough so that everyone interested in my work was able to get one and also made sure that I kept some books myself, in case someone wants one in the future. In other (managerial) words, I ordered enough to fulfill the current demand *and* I kept some inventory to face the uncertainty of future demand.

This –admittedly rather trivial example– illustrates several key aspects of inventory research. It shows, through the fundamental question, *how much, exactly, should I order?*, how inventories and orders are intimately related: inventory policies drive orders and orders drive inventories. It also demonstrates its enormous scope: not only is that question 'everywhere', but it also is repeated several times over the life of a product before it even reaches your hand. Finally, it highlights something that we (at least *I*) often forget: behind every decision, there is a decision maker with a particular set of priorities and objectives. To be able to print the number of dissertations I requested[1], the printer had to –

[1] 150, by the way.

at some point– make his own purchase decisions regarding the paper and the ink and the glue. In the same way, the paper producer (and the ink producer, and the glue producer) had to make similar decisions. Decisions that ultimately culminated in someone chopping trees to produce paper, someone extracting oil to produce glue, and someone extracting certain minerals from the earth to produce ink. The way in which trees (and oil and minerals) gradually became this book, through successive transformations (steps), and, especially, the way in which these successive steps are intimately related (decisions made at each step affect prior and later steps) is why we talk about supply *chains*.
Broadly speaking, in this dissertation we look at different ways in which inventory decisions at one point of the supply chain affect orders further up the chain. In particular, we examine different manifestations of *The Bullwhip Effect* —a phenomenon that has been studied for decades, and which explains the observations made that (1) the variability of orders tends to be larger than the variability of demand, and (2) that this amplification of demand variance itself increases the further upstream a company is in a supply chain. In other words, *The Bullwhip Effect* states that my sudden order of 150 books is capable of causing a ripple that is amplified as it travels upstream through the supply chain; by the time it reaches the lumberjack upstream, the change that he observes in his demand is far larger than the equivalent amount of trees needed to produce the 150 books I ordered. In itself, my modest order of 150 books is but a drop in the ocean, but if you consider that more than 500 million books were sold in the US alone during 2013[2], that's a lot of trees.

In this chapter, we briefly introduce *The Bullwhip Effect*; then formulate a series of research questions; introduce the methodologies used to answer these questions; and position this work among the broader operations management (OM) literature. Finally, we present a brief overview of the organization of the remainder of the dissertation.

1.1 The Bullwhip Effect

Lee et al. (1997a) define the **Bullwhip Effect** as

> "the phenomenon where orders to the supplier tend to have larger variance than sales to the buyer (i.e., demand distortion), and the distortion propagates upstream in an amplified form (i.e., variance amplification)."

[2]Source: Nielsen Bookscan.

The dynamics associated with the Bullwhip Effect have, however, been of interest for decades. For example, Procter & Gamble, is often cited (Geary et al., 2006) as the canonical example of the Bullwhip Effect in action due to their early recognition of a wave-like pattern in their diaper sales—how could pampers' demand be so profoundly cyclical, when babies' needs are, on the aggregate, relatively constant?. Forrester (1958) answered this question. Using system dynamics simulations to study amplification on supply chains from the perspective of managerial decision making, he shows that ordering decisions can drive demand amplification and create artificial seasonal effects. His contribution to the field cannot be understated; before Lee et al. (1997a) popularized its current designation, the Bullwhip Effect was known as the *Forrester effect*.

Today, we distinguish between operational and behavioral causes for the Bullwhip Effect. Lee et al. (1997a) identify four operational causes for the Bullwhip Effect: demand signal processing, rationing games, order batching, and price fluctuations. They show that the presence of any one of them is enough to induce amplification, and identify counter measures that can be applied to reduce their impact. Croson and Donohue (2006) perform an experiment based on Sterman's Beer Game simulation (1989) in which they control for all the operational causes, yet still find evidence of the bullwhip. They conclude that human behavior and the inherent difficulty of making decisions in dynamic environments contribute to the effect—in particular, the players' tendency to underestimate the pipeline inventory (orders that have been placed but not yet received) is found to be a key driver.

For a phenomenon so well understood, both analytically and in the lab, the Bullwhip Effect is still controversial and a source of inspiration to researchers all over the world. Empirical studies on the real-world appearance of the Bullwhip Effect do not draw the same clear cut conclusions as the theoretical and experimental work. Even though the Bullwhip Effect itself has been proven to be significant at the firm level (Metters, 1997; Fransoo and Wouters, 2000; Bray and Mendelson, 2012), attempts to empirically quantify the effect at higher aggregation levels have had disparate results: Studies have failed to consistently prove it statistically significant at the level of whole industries (Cachon et al., 2007; Bu et al., 2011). This lack of clear empirical evidence is attributed to the influence of factors such as the high level of aggregation (Chen and Lee, 2012), and the seasonal adjustment present in government statistics (Gorman and Brannon, 2000).

1.2 Research Objectives

The objective of the research presented in this dissertation is to obtain insights on the role of inventories in real world decision-making processes. We look at inventories through two diametrically opposed perspectives and, in doing so, answer two main research questions: *To what extent is real world decision making being driven by inventories?* And, in turn, *To what extent does the decision-making behavior affect inventory performance?*

To tackle this research objective we make the problem manageable by posing a series of smaller, more modest questions that we answer progressively through a series of studies The Bullwhip Effect is central in this dissertation; we frame all of our research questions around it, its appearance, and its link to inventory decisions and behavior.

1.2.1 Historical link of inventories and the Bullwhip

As mentioned before, research on the Bullwhip Effect in the context of supply chain management has been conducted extensively since at least Forrester's (1958) seminal work. The phenomenon is, however, so encompassing that it is of interest not only at the supply chain management level; economists have long debated about the appearance and causes of periodicity in the economy (Geary et al., 2006). Indeed, the role of inventories in the economy is a stirring topic of research, and disagreements, for micro and macro economists (Fitzgerald, 1997); the former tend to see inventories as a stabilizing force, whereas the latter tend to see them as a destabilizing force (Blinder and Maccini, 1991). In light of this, the first research question that we formulate is a meta question on past research: *What have researchers learned about the link between inventories and the Bullwhip Effect? What are the questions that are still open?* We answer this in Chapter 2, through an exhaustive literature review on the Bullwhip Effect, its history, and the link between business cycles and inventories. One interesting observation unearthed in this chapter is that research investigating the role of inventories in the general economy tends to follow cycles; interest in inventories tends to peak after economic crises and dwindle in times of boon.

1.2.2 The Bullwhip during the credit crisis

In line with the above observation, our next research question concerns the role of inventories during the recent credit crisis. Following the collapse of the financial system in September 2008, manufacturers in several industries observed a sizable drop in demand that was not driven by a comparable drop in end market demands. While end markets appeared relatively stable, the demand oscillations observed by manufacturers were massive and consistent with the evolution of demand that one would expect in a supply chain struck by a sudden shock, propagated by the Bullwhip Effect. With a demand shock out of the question, an inventory shock could provide an explanation. Thus, we formulate the following research question: *Can a synchronized inventory shock –caused by the desire of firms to retain liquidity in moments of financial distress– explain the demand dynamics experienced by upstream manufacturing firms following the collapse of Lehman Brothers on September 2008?*

To answer this question, in Chapter 3, we extend the system dynamics supply chain model from Sterman (1989) by explicitly modeling de-stocking decisions. We use primary empirical data from a Dutch chemical company to parameterize and validate models for 4 different supply chains. Methodologically, the system dynamics models draw from the behavioral operations literature: instead of assuming an optimal inventory policy, the orders for each of the firms of the supply chain (echelons) are calculated through an anchor and adjustment heuristic (Kahneman et al., 1982) calibrated with the empirical sales data. We use the crisis period as a natural experiment; the synchronization of the observed reactions allowing us to use this heuristic, previously used to explain human behavior at the individual level in controlled experiments (Sterman, 1989; Croson and Donohue, 2006).

In this work, we observe that the behavior at an echelon level closely replicates what has been reported in experiments, with firms steering on on-hand and pipeline inventories—displaying a smoothing of orders and under estimation of the pipeline. To test the validity of our hypothesis, that the inventory reduction drove the fall in orders observed by manufacturers, we construct an alternative model with no de-stocking. We find that said model is unable to replicate the observed behavior.

1.2.3 The Bullwhip and human behavior

Having found support for the capability of the anchoring and adjustment heuristic to replicate behavior at an aggregate level, we turn to its inner workings. The research question that we formulate is a descriptive one: *How does human behavior –as measured by the inventory and pipeline smoothing– affect the stability of a production/inventory system, its dynamic performance, and the amplification (bullwhip) of orders and inventories?*

We answer this question through an analytical study in Chapter 4. In this chapter, we switch our viewpoint: instead of studying the response of the system to an inventory shock, we analyze how human behavior affects inventories. To do so, we develop a model that uses a decision-making mechanism equivalent to the one used in Chapter 3 and analyze it through control theory. We base our model on the rich literature dealing with control theoretic models of supply chains (Dejonckheere et al., 2003; Hoberg et al., 2007b; Disney, 2008), however, we approach our research in a descriptive manner. If the focus of prior research in the area was primarily the development of optimal policies and guidelines for the selection of parameters (Towill et al., 2007; Disney et al., 2006a), we are interested in the effect of non-optimal behavior: Knowing that decision-makers tend to smooth orders and under-estimate the pipeline, *how does this affect the performance of the firm?*

We discover a complex interplay between human behavior, as measured by the inventory and pipeline smoothing, and the performance of the system. Non-optimal behavior, such as the under estimation of the pipeline, adds a new dimension to the trade-off between transient and stationary performance; in particular, the under estimation of the pipeline (observed by researchers in the lab and in our supply chain models of Chapter 3) introduces a cyclical component to the response of the system. In the case of inventory performance, we observe that these oscillations can result in an improved performance at the cost of an increased sensitivity of the system.

1.2.4 *The Bullwhip and seasonality*

The findings of Chapter 4 are illuminating and puzzling at the same time. By under estimating the pipeline, decision makers induce a cyclicality in the inventory behavior that –for a limited range of parameters– can potentially result in an increased inventory performance. The performance metrics used in that chapter, however, assume very specific demand conditions: The amount of bullwhip is measured against a normally distributed demand, and the dynamic performance is measured as a result of a single demand shock. Even though the reasoning behind the usage of these metrics is sound (a system that responds well to both shocks and random demand is indeed a desirable system) it leaves open the space to question its robustness. The research questions that follows are then: *How robust are the theoretically developed metrics to changes in demand?* And, in particular, *how does the cyclicality of the system interact with cyclical demands?*

We set out to answer these in Chapter 5 through a series of extensive numerical experiments carried out on the control theoretic model presented in the prior Chapter. In doing so, we turn our attention to a 'forgotten' aspect of the Forrester effect: Rogue seasonality. Rogue seasonality (originally called the 'fake business cycle' by Forrester (1958)) describes the way in which a system imposes an endogenous cyclicality to its outputs. In the case of the system we are describing, rogue seasonality describes a seasonality in the orders and inventory of the system that is different to the seasonality of demand. In the case of the theoretical, non-seasonal, demands that we study in Chapter 4, rogue seasonality is detected by the appearance of *any* kind of cyclicality in the evolution of orders and inventory. However, *how does the system react to actual demand streams?* To answer this, we first benchmark the theoretical bullwhip performance metric and find that even when the input demand is sampled from a normal distribution, the system's performance deviates from the theoretical performance. This deviation depends on the behavioral and structural parameters of the system; deviations from the optimal are heavily penalized. We find that this deviation from theoretical performance is related to the appearance of rogue seasonality and develop a method to quantify it. In the last section of this study, we analyze the interplay between exogenous (demand side) and endogenous (rogue) seasonality: The seasonality imposed by the system can attenuate the demand seasonality, provided that the frequencies are sufficiently different. With this study, we highlight the need for decision-makers to understand both their system and the demand that they are facing; optimal decision making must take both, including their interaction, into account.

1.2.5 *The Bullwhip and the chicken and the egg*

The last study in this dissertation shifts the viewpoint once again. In Chapter 5, we explicitly look at the role of inventories as drivers of decision making through an empirical study using firm-level secondary data. We construct a database of supplier-customer pairs and study how downstream inventory decisions affect upstream purchase decisions. We motivate this study through the following question: *Do upstream decision makers over-react to changes in downstream inventory levels?* We find statistical significance of manufacturers overreacting to downstream inventory changes. Furthermore, we study the origin of said inventory changes: *Can we observe evidence of firms adapting their inventory levels to economic and financial conditions?* We find evidence of this by studying a structural model of the decision-making mechanism: The data is consistent with rational decision-makers that vary their cost structure according to the economic and financial conjuncture. This finding is consistent with the hypothesis developed in Chapter 3, in which we proposed that decision makers all over the world turned to inventories as a source of liquidity—leading to a generalized de-stocking.

By focusing on answering the above set of questions, the research presented in Chapters 2–5 attempts to shed light into the role of inventories in decision-making processes. We find support for the hypothesis that explicit inventory decisions helped drive the demand drops associated with the recent credit crisis. First, we find that the behavioral mechanisms behind firm-level decision making are consistent with the individual-level mechanisms studied in the lab. Second, we identify the performance trade-offs associated with such behavior—both in a theoretical and practical level. And finally, we find that downstream inventory decisions affect upstream production orders, and that in turn, financial conditions affect these inventory decisions.

1.3 Research Methodology and Positioning

The research questions presented in this dissertation require the use of various research methodologies. In this section, we briefly describe each of them and position our work within the existing literature.

1.3.1 System Dynamics simulations

System Dynamics is popular when it comes to modeling the dynamics of complex systems, such as the production/inventory system used in Chapter 3. This methodology allows the modeler to decouple endogenous, exogenous, and structural effects. System Dynamics models explicitly describe the behavior of individual components pursuing local results, and exploit the structure of the system to model the interactions between these components.

The entire field of System Dynamics emerged from the work of Forrester (1958). In that work, the author presents an inventory/production system , with which he illustrates the concept of the dynamics we know today as the Bullwhip Effect. This model is expanded and refined by Sterman (2000) and used by other researchers to derive insights on the Bullwhip Effect (Kim and Springer, 2008). System dynamic models make two important assumptions: Continuous time and continuous flows. Because of this, the power of the methodology resides in its ability to model complex feedback-ridden systems that are able to replicate real-world behavior with relative computational ease, rather than its use in axiomatic research. In line with this, in Chapter 3 we develop an extension of Sterman's (2000) echelon model with which we run simulations that we validate empirically. The main contributions of this dissertation to the literature are the empirical testing and validation of the de-stocking hypothesis in the context of the credit crisis, and the use of this insight for managerial decision making at the tactical level.

1.3.2 Linear control theory

Linear control theory is widely used in operations management to perform axiomatic research on inventory/production systems based upon the same concepts as the System Dynamics models presented above. Control theory, in contrast to system dynamics, typically relies on transforming the mathematical time-domain representation of the model (differential equations) into an equivalent representation in the frequency domain. This is achieved through Laplace transforms (in the case of continuous time systems) or Z-transforms (in the case of discrete-time systems) (Wunsch, 1983). In the OM literature, control theoretic inventory/production systems have been studied with several different objectives. Disney and Towill (2006) show that there exists a range of policies (when the inventory and pipeline are taken into account with the same weight at the time of calculating replenishment orders) that are optimal from the perspective of stationary and transient performance, these policies are called Deziel-Eilon (DE) policies in honor of the researchers that first described them (Deziel and Eilon, 1967); Dejonckheere et al. (2003) shows that when the adjustments of DE-policies are full (inventory and pipeline deviations are taken entirely into account), then they are equivalent to regular Order-Up-To policies, and introduce the concept of fractional adjustments as a tool to reduce the bullwhip. Disney (2008) explores non-optimal policies and presents a way of calculating the stability of the policies for a given lead time through a determinant-method based on Jury's (1964) inner method. In Chapter 4, we develop a compact expression for the stability of the system for any arbitrary value of lead time. In Chapter 5, we perform numerical experiments on the control theoretic models to characterize the appearance of rogue seasonality in the system's orders.

1.3.3 Empirical modeling

There exists a wide range of inventory research that explicitly connects several stages in the supply chain (Graves, 1999). We use firm level, secondary empirical data to test econometric forecasting and inventory models at two stages of the supply chain. The contribution of our research to the OM literature is the inclusion of two stages of the supply chain into these models. The second part of Chapter 6 presents a structural model of the processes behind the inventory decision making that takes economic and financial conjuncture into account to decide upon the cost-ratio that the firm uses for determining the optimal value of its safety stock. In structural modeling, theory is used to develop mathematical statements about how a set of observable "endogenous" variables y, are related to another set of observable "explanatory" variables, x. In this context, we assume that decision makers are rational, and then use the data to impute the cost parameters that explain the observed behavior. In the OM literature, structural modeling has been used most notably by Olivares et al. (2008), who used a similar approach to impute the implied costs behind the decisions on operating room assignments.

1.4 Outline of the Thesis

The rest of this dissertation is divided as follows: In Chapter 2, we present a literature review that focuses on the historical research conducted on the Bullwhip Effect, and the role of inventories within said literature. Chapter 3 follows with the empirical modeling of supply chains during the credit crisis. In this chapter, we develop a de-stocking hypothesis and test it with primary empirical data; we show that upstream companies experienced market dynamics consistent with a synchronized de-stocking shock affecting the world's supply chains, and study the underlying managerial behavior from a behavioral operations point of view. In Chapters 4 and 5, we use a discrete time control theory model of the inventory/production system to investigate the role of human behavior on inventory and order performance. The former is an analytical study, where we provide a general expression for the stability of the system under any kind of behavioral policy, and the latter is a numerical study, where we perform extensive experimentation to understand the effect that different demand patterns have on the performance of the system. Chapter 6 returns to the empirical world, to test the hypothesis that downstream inventory decisions affect –and cause overreactions– upstream production decisions. We use secondary data from customer-supplier pairs to test econometric and structural models.

> Understanding the world requires you to take a certain distance from it. Things that are too small to see with the naked eye, such as molecules and atoms, we magnify. Things that are too large, such as cloud formations, river deltas, constellations, we reduce. At length we bring it within the scope of our senses and we stabilize it with fixer. When it has been fixed we call it knowledge.
>
> Karl Ove Knausgård

Chapter 2

Literature Review

In the introduction, we briefly described the Bullwhip Effect and the different methodologies used to study it. In this chapter, we review the existing literature on the Bullwhip Effect by positioning published contributions to the theory in a different way; rather than grouping different works according to their *methodology*, we group them according to their *purpose*.

We first introduce the stream of research that attempts to *measure* the Bullwhip Effect. This includes the pioneering work that predates the term "Bullwhip Effect", but sets the mathematical basis for everything that came after; the work that sets performance measures for different aspects of the bullwhip; and the empirical work that seeks for evidence of the Bullwhip Effect in real-life data.

Then, we turn our attention to the *analysis* of the bullwhip: How different replenishment policies and forecasting methods impact the bullwhip; how human behavior exacerbates the information distortion; and how frequency domain analysis can help uncover a hidden dimension of the bullwhip: rogue seasonality.

Finally, we take a look at the literature that proposes and analyses different ways of *taming* the bullwhip, primarily through the usage of particular policy settings and/or information sharing.

2.1 Discovering and measuring the Bullwhip Effect

Before we discuss the Bullwhip Effect, we must introduce the mathematical structures that made it possible for researchers to identify the phenomenon. Forrester is considered the "father" of the Bullwhip Effect due to his seminal 1958 work. However, his discovery of the Bullwhip Effect as a systematic phenomenon could not have happened had he not, first, developed the mathematical tools of system dynamics. These tools allowed him to numerically solve complex inventory models based upon difference and differential equations.

In parallel to Forrester's progress, researchers like Simon (1952) and Vassian (1955) started to use transform methods to obtain analytical solutions to these models. Lacking the computing power we have today and years of refinement of these methods, the analytical study of dynamic models would trail the numerical study of complex system dynamic models for decades.

In this section, we first introduce the early era of dynamic inventory models, pioneered by Simon's (1952) and Vassian's (1955) usage of transform methods. Then, we follow the path that took us from these mathematical formulations to the control theoretic inventory modeling framework, which adapts, and applies, concepts originally developed in systems engineering to the analysis of dynamic inventory models. Following this, we introduce Forrester's (1958) work and its legacy: the development of system dynamics and the discovery of the bullwhip. Having established the existence of the bullwhip, we then present the efforts of researchers determined to introduce appropriate ways to quantify it; and we finish by recounting the past, and present, controversies and difficulties inherent to the search for empirical evidence of the bullwhip.

2.1.1 Before the bullwhip: Mathematical foundations of dynamic analysis

Simon's (1952) exploratory study laid down the mathematical foundation for the analysis of the dynamics of production control. In his paper, the author points out the 'obvious analogies' between electrical and mechanical control systems, used to control physical systems through the use of servomechanisms; and production control systems, used to plan and schedule production in a business setting. Simon develops a simple model of a production control system in terms of servomechanisms operating in continuous time. He takes advantage of the recent (at the time) advances in techniques and methodologies developed for engineering applications, to analyze the dynamic response of this model. Simon recognizes that the (inherently dynamic) variation in inventories and orders drive the total cost of the system and, using Laplace transform methods, analyses a quadratic cost system. He finds that optimizing for steady state will not generally minimize cost for changes in the steady state. Thus, he defines a decision rule for the rate of change of production that minimizes the steady state costs while dampening transient deviations. Furthermore, he shows that stabilizing inventories leads to fluctuating orders and vice versa. The insights obtained in this study are limited in their application and – by the author's admission– readily obtained by intuition; however, in setting up the first dynamic model of an inventory system, this study immediately opens the door for more. Simon's work is the basis for two methodological branches of the utmost importance for the study of supply chain dynamics and the Bullwhip Effect: The control theory methodology, which uses Laplace and Z transforms to analytically solve increasingly more realistic and complex systems; and the system dynamics field, which essentially allows for a fast numerical solution of the represented systems and thus permits the modeling of even more complex systems. Extending from this seminal work, Vassian (1955) will lead the analytical branch by developing discrete time models (solved using the Z transform), and Forrester (1958) will develop system dynamics.

2.1.2 *Control theory and the mathematical analysis of production systems*

Vassian (1955) extended Simon's (1952) work by describing an equivalent production system using discrete time. For this discrete time, periodic-review system, with deterministic lead times and complete back-ordering, the author defines the following difference equation:

$$I_k - I_{k-1} = \theta_{k-(T+1)} - C_k, \tag{2.1}$$

where I_k represents the inventory at time k, θ_k is the order quantity placed at time k, T represents the production lead time, and C_k is the customer's order quantity at time k. Vassian recognizes that a desirable decision rule depends on both the forecast of demand and the inventory level at each time period. Specifically, he finds that the following decision rule:

$$\theta_k = C_k^*(T+1) - \sum_{j=1}^{T} \theta_{k-j} - (I_k - I_p), \tag{2.2}$$

where $C_k^*(T+1)$ is the forecast of total customer orders during periods $(k+1)$ until $(k+1+T)$, results in the minimum variance of inventories about the desired level, for any given sequence of forecast errors. In particular, for normally distributed errors, the mean inventory level, I_p, can be easily determined to give any arbitrary service level. The author's analysis, however, makes no explicit consideration for either a cost structure or the variability of production. Deziel and Eilon (1967) tackle the problem of the interplay between inventory and production costs in a paper that will prove to be an important stepping stone in the control-theoretic supply chain literature. Even though it effectively minimizes stock-holding costs by minimizing inventory variance, they argue that the above decision rule is inadequate when considering total costs as a combination of stock-holding, shortage, and production fluctuation costs. They construct a variant of the decision rule by inserting a smoothing constant γ that allows for production (or order quantity) fluctuations to be reduced at the expense of increased inventory variability[1]:

$$\theta_k = \gamma \left[I_p - I_k - \sum_{j=k-T}^{k-1} \left(\theta_j - F_j \right) \right] + F_k, \tag{2.3}$$

with F_k a simple exponentially smoothed forecast of the demand at time k. Deziel and Eilon show that with the above decision rule, the system is stable for values of γ and α (the exponential smoothing parameter) between 0 and 2. The authors then conduct a series of simulations with an analogue computer to characterize the behavior of the system to varying values of γ, α, and T. Additionally, they define (and compute) a measure of the system's performance as the ratio between demand variance and order (and inventory) variance when the demand is stationary and stochastic. This performance measure is equivalent to the measurement of the bullwhip later adopted by Fransoo and Wouters (2000). The mathematical tractability of this decision rule, coupled with its stability and performance make it the most adopted rule in the control theoretic inventory modeling field. Decades after its publication, this decision rule will be the object of intense study by researchers in the field—it has been retrospectively been named the DE-APVIOBPCS design.

Towill (1982), formalized the concepts described here with the introduction of the Inventory and Order Based Production Control System (IOBPCS) design framework. In an IOBPCS design, replenishment orders are generated as the sum of an exponentially smoothed demand forecast and a fraction of the inventory discrepancy. In contrast with Deziel and Eilon's (1967) decision rule, the pipeline is not involved in the calculation of replenishment orders for

[1] For consistency, all the control theoretic models are presented, in this chapter, with the notation used by Vassian (1955).

IOBPCS models. However, by representing an Inventory/Production system in block diagram form, the work of Towill allowed for the straightforward application of linear control theory methodologies to study structural and dynamic properties and sparked numerous extensions to the models. Given that a substantial amount of the extensions to the IOBPCS models concern the choice of smoothing parameters (γ), we present a graphical sketch of the subsequent extensions as a function of inventory and pipeline smoothing parameters (γ_I and γ_P). This is shown, together with the classical Order-Up-To (OUT) policy for reference, in Figure 2.1

A first extension to IOBPCS is VIOBPCS (Variable Inventory and Order Based Production Control System), where the inventory target is no longer constant, but calculated each period as a multiple of the demand forecast. Edghill and Towill (1990) study this system and find that, in comparison to IOBPCS, the variable inventory targets of VIOBPCS designs introduce interesting trade-offs between the "marketing" and "production" sides of a firm: increased service levels through a better correlation of inventory and demand, at the cost of increased variability in orders.

A powerful extension, APIOBPCS (Automatic Pipeline Inventory and Order Based Production Control System) developed by John et al. (1994), adds a second feedback loop in the form of a pipeline adjustment. APIOBPCS generalizes the decision rule of Deziel and Eilon (1967) by allowing independent smoothing of the pipeline and inventories. Adopting the notation used above, we write the orders generated by an APIOBPCS [system]:

$$\theta_k = \gamma_I \left[I_p - I_k \right] + \gamma_P \left[\sum_{j=k-T}^{k-1} \left(F_j - \theta_j \right) \right] + F_k, \tag{2.4}$$

where γ_I and γ_P are the independent smoothing parameters for inventory and pipeline respectively. A final extension to this system; APVIOBPCS (Automatic Pipeline Variable Inventory and Order Based Production Control System) drops the assumption of a constant desired inventory level, I_p, and replaces it by a function of the expected sales, thus:

$$\theta_k = \gamma_I \left[CF_k - I_k \right] + \gamma_P \left[\sum_{j=k-T}^{k-1} \left(F_j - \theta_j \right) \right] + F_k, \tag{2.5}$$

where the coverage C is constant, but the resulting desired inventory varies with the updating of the forecast. Dejonckheere et al. (2003) show that this decision rule is indeed a modification of an Order-Up-To policy with a safety lead time in place of an explicit safety factor multiplying the demand variability[2]. Despite the equivalence of APVIOBPCS and OUT policies, the control theory literature focuses on the analysis of the influence of parameters and the structure of the system in the dynamic response of the system, often without explicit considerations of cost functions. We review the control theoretic literature's role on analyzing the dynamic performance of supply chains in section 2.2.

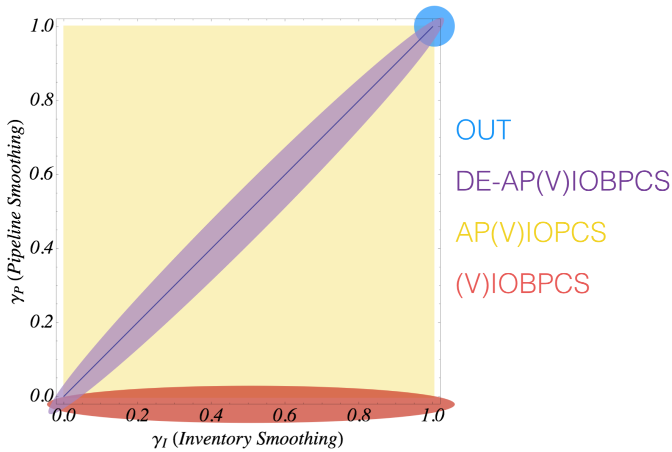


Figure 2.1 *Overview of IOBPCS design variants according to pipeline and inventory smoothing approaches.*

2.1.3 *System dynamics and the Forrester effect*

According to Forrester (1958), system dynamics (originally dubbed industrial dynamics) was created as a tool to aid managerial decision making. In this view, his objectives and approach to model design are pragmatic; insights come first, mathematical rigorousness, second:

> Industrial dynamics is an approach that should help in important top–management problems. (...) Many men predetermine mediocre results by setting initial goals too low. The attitude must be one of enterprise design. The expectation should be for major improvement in the systems. The attitude that the goal is to explain behavior, which is fairly common in academic circles, is not sufficient. The goal

[2]Consider a classical OUT policy, where orders at time t, O_t, are calculated through $O_t = S_t - \text{Inventory Position}_t$, with $S_t = D_t^L + k\sigma_t^L$; D_t^L the expected lead time demand, σ_t^L the standard deviation of the lead time demand, and k a safety factor. If we set up $k = 0$ and increase the coverage lead time by a safety factor C, then the policy defined by Equation (2.5) is a complete analogue to the OUT policy when $\gamma_I = \gamma_P = 1$, and a smoothed extension of the OUT otherwise

> should be to find management policies and organizational structures that lead to greater success. (Forrester, 1958).

Forrester proposed a substantial paradigm shift, so much so that his book ends with an appendix dedicated to teaching his new concepts to beginners. He stresses the need for designers to ask the right questions, to think about the correct scope of the system to be modeled, to think in terms of feedback structures, and not to be afraid to think of models as ever evolving—in his view, as the understanding of the system by the designer evolves, so should the model. It comes as no surprise, then, that the introduction of this new methodology was not without questioning. Ansoff and Slevin (1968) attempt an impartial discussion on the merits and shortcomings of system dynamics as a methodology; they argue that even though it can provide "*a useful* way of looking at a business firm", it can not be said to be "*the way* of looking at a business firm". In this view, for all its merits, system dynamics falls short of a proven theory; the approach is considered suitable for modeling complex systems, but it contains insufficient support when it comes to analysis (Ortega and Lin, 2004).

Ansoff and Slevin (1968) cite the inventory modeling of Simon (1952) as a successful attempt to translate the mathematical feedback control theory into "theorems and generalizations applicable to the firm" thus, conserving the analytic support. Simon's approach –as we have seen– sparked a vast amount of rigorous research, but –in what can be interpreted as a justification of Forrester's points– it took decades for researchers to develop analytic insights equivalent to those shown by Forrester.

Forrester's implementation of Simon's inventory model starts with a subtle mathematical inversion of the problem; rather than expressing the system as a set of differential equations, he expresses the system as a set of integral equations—he decomposes the system into rates and stocks, the latter being the integral of the former.[3] With this in mind, the continuous time integral equation for inventories is:

$$I(k) = \int_0^k [O(t) - C(t)]\, dt + I(0), \tag{2.6}$$

[3] Forrester (1958) argues that, even though engineers feel comfortable describing the world in terms of differential equations, nature "tends to integrate rather than differentiate". To illustrate this point, take the relationship between acceleration and velocity; engineers tend to see as a differential equation—with velocity being the derivative of acceleration. Forrester argues that the natural relationship is the inverse: Forces create acceleration, which is integrated to get velocity. Both views are mathematically equivalent, yet conceptually opposite.

where $I(0)$ is the initial inventory level, $O(t)$ is the rate at which inventory is received, and $C(t)$ is the consumption rate of the inventory. This integral equation represents a stock, or level. The ordering decision rule, on the other hand, is a rate equation that is itself linked to the level equation—this is the reason why we speak of feedback loops.
Analyzing equation (2.5) from this viewpoint, we distinguish one feedback loop for the pipeline correction, and one feedback loop for the inventory correction. It follows, then that in a system dynamic representation of this system, we divide the system into 2 level equations (inventory and pipeline) and two rate equations (inventory and pipeline arrival rates), each with a feedback loop. The power of system dynamics resides in splitting complex systems into smaller sub-systems that can be represented in terms or levels, rates, and feedback loops. The rate equation for the inventory subsystem is

$$O(k) = \frac{\hat{I} - I(k)}{\mathrm{AT}}, \tag{2.7}$$

where $\hat{I}$ is the desired inventory level and AT is the adjustment time.[4]

The approach taken in system dynamics to solve the system of simultaneous equations is to split the time interval into smaller discrete chunks DT, assume that rates are constant in each of these intervals and, starting with a initial values, calculate a new level by using the value of the slope (rate) at that point, and consequently calculate the new value of the rate using the new level. Figure 2.2 shows a graphical sketch of this approach (Figure adapted from Ortega and Lin (2004)). The given levels and rates are calculated sequentially through:

$$I(a) = I(0) + \mathrm{DT} \times (O(a) - C(a)) \tag{2.8}$$

$$O(b) = \frac{\hat{I} - I(a)}{\mathrm{AT}}. \tag{2.9}$$

Forrester justifies this numerical approach by putting the insights obtained ahead of rigorousness:

> A fascination with methodology is apt to lead to a quite disproportionate amount of attention to peripheral questions of technique. The mathematician may want to substitute some more elegant methodology for the first-order integration such as used in this book,

[4]The adjustment time in system dynamics models represents the time that we allow for the adjustment to take place. As before, the thinking behind this adjustment time is just a switch in the viewpoint we used in the previous decision rules. Mathematically speaking $\mathrm{AT} = \frac{1}{\gamma_I}$, thus, a long adjustment time is tantamount to a heavy smoothing of the inventory difference.

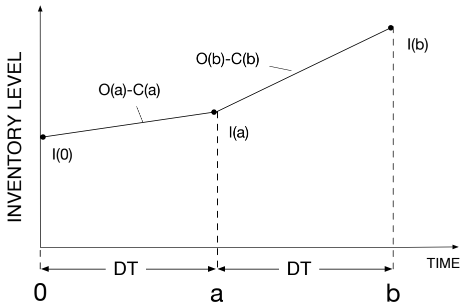


Figure 2.2 *Sketch of inventory levels and rates.*

> even though there can probably be no objective demonstration of the necessity. He may become involved in the discontinuities and the computing phenomena that revolve around the selection of a solution interval DT. He may become involved in trying to determine how large this interval can be made, rather than merely making sure that it is small enough to raise no questions. (Forrester, 1958).

Forrester uses this methodology to construct a 4 echelon supply chain model (retailer-wholesaler-distributor-factory) and test the performance of the system with various retailer demands and supply chain parameters. His simulation results show the tendency that orders at a given echelon tend to be more variable than demand, that this variability increases further upstream, and that random downstream demand fluctuations can generate cyclical fluctuations in upstream demand. Forrester, through his system dynamic simulations, exposed what we know today as the Bullwhip Effect—and was for decades referred to as the Forrester effect. Burns and Sivazlian (1978) use a combination of simulation and analysis of a multi-echelon supply chain to investigate the causes of the Forrester effect, and find two different sources: the necessary adjustments due to transient inventory changes, and "false orders" generated by upstream echelons misreading the change in downstream orders. They show –through simulations– that the first of these sources is unavoidable when echelons want to achieve a target inventory level, but that the second source is a byproduct of the way that information is transmitted and thus preventable. They develop a decision rule based upon the parameters of the system that manages to eliminate the false order effect.

Forrester's system dynamics group at MIT developed, during the early 60's, the beer distribution game as a showcase of the dynamics behind multi-echelon production/distribution systems. To this day the beer game has been played by millions of students and managers, and remains the go-to tool for teaching the mechanisms and insights behind the Bullwhip Effect. Sterman (1989) uses the causal feedback structure from system dynamics to model the beer game dynamics. Using data from beer game sessions played by students, and theory from human behavioral research developed by Tversky and Kahneman (1974), he fits a decision rule based upon (2.5) and finds that human players consistently underestimate the pipeline ($\gamma_P < \gamma_I$) when calculating replenishment orders. Under the umbrella term "misperceptions of feedback", Sterman describes the subjects's tendency to ignore past decisions, provided they have not yet taken effect, when making future decisions. This study is significant in the search of the Bullwhip Effect because it explicitly introduces human behavior to the list of causes of the demand amplification. Sterman (1989) ends his paper calling for the development of testable theories to explain the link between individual human behavior and observable macro behavior.

It is the work of Lee et al. (1997b) that opens the door for mainstream acceptance of the Bullwhip Effect. In this seminal paper, the authors acknowledge Forrester's and Sterman's contribution to the understanding of supply chain dynamics and study the appearance of the Bullwhip Effect from an analytical perspective, using traditional multi-period inventory models and order up to policy decision rules. The authors find 4 structural causes of the Bullwhip Effect: Demand signal processing, order batching, shortage gaming, and price variations. They demonstrate that each of the causes is independent of each other, and relate them to the relaxation of 4 commonly used assumptions in inventory modeling (stationary demand, infinite resupply with fixed lead time, zero fixed ordering cost, and stationary production costs). This study is pivotal not only for popularizing the name of the Bullwhip Effect[5] but because it shows that: (a) rational, optimizing, managerial behavior is capable of triggering the bullwhip, and (b) for explicitly recognizing the difference between material and informational flows, the latter being the source of distortion. We review this paper in more detail in section 2.2.

2.1.4 *Measuring the bullwhip*

Thus far in this review, we have settled on the Bullwhip Effect as the name for the phenomenon of variability amplification in the supply chain, and provided evidence that human behavior (Forrester, 1958; Sterman, 1989), as well as rational behavior (Lee et al., 1997b) are potential causes of its appearance. We focus now on the literature related to the objective *measurement* of the Bullwhip Effect.

[5]Prior to the publication of this paper, Forrester effect, whiplash effect, and whip-saw effect where also popular terms to refer to the amplification of demand

In the studies we have examined this far, Forrester (1958) measures the amplification in a supply chain by quantifying the peak overshoot to a demand shock; (Lee et al., 1997b) compare the variances of orders and demand, and claim that the bullwhip exists when the former is larger than the latter; and Deziel and Eilon (1967) –without explicit references to the Bullwhip Effect– introduce the quotient between order and demand variability under stationary random demand as a measure of the stationary performance of the system. These methodologies all quantify the bullwhip, however, they produce measures that are not comparable.

Fransoo and Wouters (2000) tackle the problem of measuring the bullwhip from the perspective of industrial practice. The authors discuss the implications of data aggregation; they show that when presented with ample data from electronic point of sales (EPOS) data from multiple products and multiple firms, different aggregation strategies result in different measures of the bullwhip. The authors adopt the quotient of the coefficient of variation of orders and the coefficient of variation of demand as a standard measure for the bullwhip and recommend that the bullwhip be measured for each echelon in the supply chain independently and, in case of the existence of disaggregated data, its aggregation be based upon the specific problem under investigation. Disney and Towill (2006) remark that in addition to the mentioned measures, frequency response plots (Dejonckheere et al., 2002) are relevant for the measure of artificial seasonality in the response. Frequency response plots, however, require frequency domain analysis, which limits their application in analytical work. Dejonckheere et al. (2003) show that when the input is an i.i.d normally distributed demand, a common metric in communications engineering (Garnell and East, 1977), the noise bandwidth, is proportional to the bullwhip as measured by the quotient of variances. Disney and Towill (2003) additionally remark that to properly quantify the bullwhip of a system, the quotient between the variance of inventory and the variance of demand must also be considered (recall that stabilizing orders tends to increase the fluctuations of inventories and vice versa). Quantifying the Bullwhip Effect through these variability quotients (for simplicity we refer to these measures as BW_O, or bullwhip of orders, and BW_I, or bullwhip of inventories) has an attractive quality: it can be measured empirically as well as calculated theoretically (provided we make assumptions on the demand distribution); however, it is a static measure—it gives us information of the overall response of the system but ignores the transient performance. Hoberg et al. (2007b) analyses different transient performance measures and concludes that the Integral Time-weighted Absolute Error (ITAE) is an appropriate measure of the system's performance. The ITAE accumulates the time-weighted deviations from the ideal performance of the system after it is presented with a single unit demand shock; it quantifies the magnitude of the shock amplification by accumulating the errors, and it penalizes long lasting deviations by time-weighting them.

One dimension of the Bullwhip Effect that has been proven difficult to quantify is the rogue seasonality, namely the appearance of a seasonality in the orders and inventories that is not present in the demand. Metters (1997) shows that eliminating demand seasonality brings substantial cost savings and links his analysis to the Bullwhip Effect by citing it as an important source of seasonality. His analysis, however, does not attempt to either provide a solution or a measurement of rogue seasonality. Recent work, inspired in the control of chemical process plants by Thornhill and Naim (2006) introduces Spectral Principal Component Analysis (SPCA) as an empirical measure of endogenous seasonality, but this measure is qualitative and thus of a limited application. Kim and Springer (2008) perform a theoretical analysis of the conditions under which rogue seasonality appears in a system dynamics supply chain model; they define a rogue seasonality measure and suggest strategies to minimize volatility, however, their analysis approximates lead times through a first order exponential delay. Shukla et al. (2012) extend the empirical measure of rogue seasonality and quantify its appearance on a system dynamics simulation model.

2.1.5 The empirical conundrum

Anecdotal evidence of the Bullwhip Effect is often presented as motivation for the study of the bullwhip: P&G famously detected (and later mitigated) the presence of strong seasonal patterns in their diaper products—one would expect the demand for *pampers* would be fairly stable! (Lee et al., 1997a); Barilla has also faced a similar problem in their pasta supply chain (Hammond, 1994); and HP in their printer consumables (Lee et al., 1997a). But the plural of anecdote is not data. Even though we understand the theoretical and behavioral causes of the Bullwhip Effect –and can replicate it in the lab– formal empirical research is needed if one is to conclude that it is indeed present in the real world.

Using the bullwhip measures to quantify the bullwhip in an empirical setting is, in theory, straightforward: given demand, order, and inventory time series we can calculate the bullwhip of any system. In practice, however, we face the reality of data availability. Data is often aggregated at arbitrary levels (in terms of time and product); data is often unavailable, in such cases it must either be estimated or a suitable proxy needs to be found. To illustrate the importance of data aggregation, Fransoo and Wouters (2000) perform an empirical study of the Bullwhip Effect exploiting the availability of fine grained data for multiple products and outlets. The authors use their data to calculate the bullwhip using 4 different types of data aggregation: individual product and outlet, individual product and combined outlet, individual outlet and combined product, and combined product and outlet; they find that the aggregation strategy has a significant impact in the measured bullwhip, and consequently recommend that researchers base the sequence of data aggregation on the specific problem that is being investigated. Unfortunately, as researchers, we must often use secondary data over which we have no control—different databases impose different trade-offs in their granularity. For example, COMPUSTAT offers firm-level data of American companies aggregated at quarterly intervals, the United States Census Bureau offers industry level data aggregated at the monthly level, and EUROSTAT provides industry level data for 27 European countries in monthly intervals, but lack inventory data.

As the title of this section suggests, empirical evidence of the Bullwhip Effect is disparate: empirical evidence of amplification is found in some studies, empirical evidence of smoothing is found in others. Given this, researchers have dedicated their time to understand these results; is the diversity caused by the fact that ultimately different firms perform differently? Or is the diversity a by-product of data aggregation and manipulation? In general, the more disaggregated the data used in a study, the stronger the evidence of the bullwhip. This is in line with Fransoo and Wouters's (2000) results mentioned above, and also with the results of a recent study by Chen and Lee (2012).[6] At an SKU level, Trapero et al. (2014) find evidence of bullwhip in 100% of a dataset comprising of a year of weekly sales and incoming shipment observations of 16 SKU's sold at a UK retailer. Lai (2005) presents a study based on a larger sample; with a database of 3745 SKU's from a Spanish supermarket retailer, collected with a monthly frequency over 28 months, his analysis finds evidence of amplification in 80% of his sample. Moreover, he studies the causes of the amplification and concludes that the main contributor to the bullwhip is order batching. At a higher aggregation level, Shen (2008) conducts a study using product-category-subgroups data published by the Department of Statistics of the Ministry of Economic Affairs in Taiwan. The author constructs a database of 10 years of monthly observations of production, sales, and inventories for 15 key group categories of the electronic part manufacturing sector (such as integrated circuits, LCD's, and PCB's). He finds that only 3 out of the 15 groups exhibit bullwhip[7]. Additionally, the author investigates the effect of temporal

[6] In fairness, one could suggest that a positive bias could exist in individual level studies and products that do not exhibit the Bullwhip Effect would simply not be subjects of study.

[7] One could argue that these product categories have strong incentives to have a smooth

aggregation in his data, and finds that performing the same tests with data aggregated over quarters changes the value of the bullwhip measure but not enough to alter his insights.

Bray and Mendelson (2012) use quarterly COMPUSTAT data to estimate the bullwhip of 4297 public U.S. companies over the period 1974-2008; they find that about 66% of the firms in the sample exhibit a bullwhip. Additionally, they also find that comparing the period 1974-1994 to 1995-2008, the sample's mean bullwhip dropped by 33% and hypothesize that this drop can be related to advancements in information systems. In an empirical study using data from Chinese manufacturing industries, Bu et al. (2011) use quarterly data for the period 2004-2010 and finds support for bullwhip in 23 out of the 27 industries in the sample. Dooley et al. (2010) use data from the U.S. Bureau of Economic Analysis to investigate the effect of the recent financial crisis on the American industry. Using data aggregated at retail, wholesaler, and manufacturer levels the authors find empirical support for the hypothesis that U.S. manufacturing firms observed a Bullwhip Effect during the recession. Finally, Cachon et al. (2007) perform an exhaustive analysis of U.S. industries using monthly Census data aggregated at the 3-to-5-digit NAICS level for the 1992-2005 period. They find support for the bullwhip in 17% percent of retailers, 89% of wholesalers, and 40% of manufacturers. Recognizing the contrast of their results with previous studies, the authors conjecture that their usage of seasonally unadjusted data could be a potential explanation for the low proportion of bullwhipping manufacturers. They hypothesize that the usage of seasonally adjusted data biases the data of other studies towards amplification. This conjecture is consistent with the findings of Ghali (1987) who, in a prior study, did not find support for amplification when using seasonally unadjusted data of the cement industry (only 2 out of 19 Portland cement production districts exhibit a bullwhip). After seasonally adjusting the same data, the proportion of districts that exhibit a bullwhip changes from 2/19 to 8/19. Gorman and Brannon (2000) perform a similar study using monthly Census data for U.S. manufacturing aggregated at the 2-digit SIC level for the 1958-1997 period. They find that seasonally adjusted data overestimates the bullwhip in the majority of cases.

Chen and Lee (2012) tackle the potential causes of bullwhip masking in empirical data head-on; they analyze a modified base stock policy with capacity constraints and order batching—2 features that are expected in real-life systems. They find that capacity constraints dampen the amplification of orders, and show that using shipments as a surrogate for demand –like Cachon et al.

production.

(2007) and most empirical studies do– also result in an under-estimation of the bullwhip ratio; they explicitly differentiate between informational bullwhip (using demand and placed orders) and material bullwhip (using shipment and receipts). With regards to seasonality, Chen and Lee (2012) also find analytical support for the findings of Ghali (1987), Gorman and Brannon (2000), and Cachon et al. (2007): When the variability of the seasonality dominates over the variability of demand then, if the bullwhip is present, the inclusion of seasonality in the data dampens the bullwhip ratio. Additionally, Chen and Lee (2012) show that when the system has a finite capacity, the inclusion of seasonality dampens the material bullwhip ratio. In terms of aggregation, the authors show that temporal aggregation masks the bullwhip ratio. Also, that product and location aggregation can mask the bullwhip when it is subject to other conditions, such as the auto-correlation of the demand streams and shared seasonality patterns.

In a recent working paper, Chen et al. (2014) argue that a discrepancy between empirical studies and the theory exists due to common assumptions made in the former. In particular, the authors claim that while the theoretical studies of the Bullwhip Effect measure an information bullwhip (through the variance of orders and demand), empirical studies use material flow data (sales and shipments) as a proxy for the information flow data. This results in a measure of a material bullwhip, which is assumed equal to the information bullwhip. In this study, Chen et al. develop an exact characterization of material flow in a two-echelon model and find that, in general, the information and material bullwhips are not necessarily equal. Furthermore, they identify a series of factors that affect the discrepancy between the bullwhip measures: The stocking level of the supplier, supply chain location, lead time, and demand correlation. The conclusion derived from this work is that, because empirical research typically quantifies the Bullwhip Effect through material flow measurements, such studies should control for the mentioned factors so as to avoid estimation biases[8].

In summary, empirical evidence of the Bullwhip Effect is not universal, nor should it be. As Ghali (1987) points out, if firms seek to smoothen their production, the degree to which they do it depends on cost factors (e.g., how expensive it is to vary production as opposed to varying their stock levels) and the seasonality present in the particular demand. Looking back at the empirical research performed in the past decades, we see that the Bullwhip Effect is significant and consistently measurable at lower levels of aggregation. Its evidence at higher aggregation levels is not as clear cut, but we have theory that

[8]Whether the material bullwhip overestimates or underestimates the information bullwhip depends on the combination of the prior factors, and is thus, impossible to characterize the material bullwhip as being “always overestimating” or “always underestimating” the information bullwhip.

explains the reasons why this happens. Sprague and Wacker (1996) argue that perhaps better inferences could be made if the data were aggregated by echelons along the inventory stream because the business objectives behind each stage differ. Industries in Cachon et al. (2007) are aggregated at a level of between 3 and 6 digit NAICS code[9]. Using a 3-digit NAICS code aggregation, such as code 332 "Fabricated metal products" lumps together firms with different objectives, lead time, and distance from the end market[10], all factors that are potentially crucial in the appearance of the Bullwhip Effect (Hosoda and Disney, 2006).

In a recent paper, Bray and Mendelson (2013) argue that production smoothing and the Bullwhip Effect are not necessarily mutually exclusive—a firm may try to smooth its production, and yet exhibit the Bullwhip Effect. Because of this, they suggest that measuring production smoothing by comparing the variance of production to the variance of orders is a comparison between "apples and oranges". They develop a new measure for production smoothing and test it with monthly data from the automotive sector disaggregated at the car model level. They find evidence of both production smoothing and bullwhip present in the data.

2.1.6 A look at economists

The dichotomy between production smoothing and variance amplification has not been the sole territory of researchers in the Operations Management/Operations Research field; it's at the heart of a puzzle that has entertained macro economists for decades. The vast majority of macroeconomic research on inventories models is based upon production smoothing models (Fitzgerald, 1997). These models were first introduced by Holt, Modigliani, Muth, and Simon (1960) in the context of research aimed to improve decision-making at a factory-warehouse system. In its most basic form, the intuition behind the production smoothing model, is that in the face of varying demand and convex production costs, a profit maximizing firm will use inventories to buffer the sales fluctuations and thus maintain production as stable as possible (Blinder, 1986). This implies that inventories will decrease when sales increase, that inventories will decrease when sales increase, and that –because of the buffering of inventories– production will be more stable than sales.

[9]Under the NAICS classification, the first 2 digits define the economic sector, the third designates the sub-sector, the fourth the industry group, the fifth the NAICS industry, and the sixth the national industry.

[10]Because the NAICS classification is hierarchical, code 332 includes all 4-digit codes starting with 332; 3321 "forging and stamping", 3322 "cutlery and hand tool manufacturing", 3323 "architectural and structural metals manufacturing", etc.

Researchers in the field, however, agree that empirical data shows that inventories and sales are pro-cyclical (they grow together) and that the variance of production is larger than the variance of sales Wen (2005)—this, of course, goes against the main predictions of a simple production smoothing model and is a puzzle that has sparked a vast amount of research. In this section, we offer a stylized summary of the historical evolution of the view of inventory dynamics within the macroeconomic field.

Before the prevalence of production smoothing models, dynamic models dominated the field. Lundberg (1937) developed the first of such models by introducing a lag whereupon decision-makers had to base production decisions on sales information from the previous period. Metzler (1941), in a paper developed while completing his PhD in Harvard, extended this model in two key ways: (i) he analyzes the influence of sales expectations on the dynamics of the system, and (ii) he introduces the concept of a variable target inventory level, where decision-makers attempt to maintain inventories at a constant proportion of expected sales. The author shows that such systems generate endogenous inventory cycles and performs numerical experiments to characterize the behavior of the system to changes in the parameters. Because of the role of inventories in such systems is de-stabilizing, such models are known in the literature as "inventory accelerator" models. Goodwin (1948) and Lovell (1961) study the "flexible accelerator model". In this model, target inventories ($\hat{I}_t$) are assumed to depend on sales (S_t) through:

$$\hat{I}_t = a + bS_t, \tag{2.10}$$

with a and b constants. In its simplest incarnation, the flexible accelerator assumes that, because there exist costs in adjusting the inventory, the actual inventory change in the period (ΔI_t), will depend on an arbitrary fraction d:

$$\Delta I_t = d(\hat{I}_t - I_{t-1}), \tag{2.11}$$

where I_{t-1} is the actual inventory level at the end of period $t-1$, and $0 < d < 1$. This results in production releases (Q_t) depending on the period sales[11] (S_t) and the change in inventory levels:

$$Q_t = S_t + \Delta I_t. \tag{2.12}$$

In particular, Lovell uses an extension of this framework to perform an empirical

[11] The authors do not explicitly state a sequence of events. Production at time t depends then on sales during the same period and the difference between initial and ending inventory levels.

study of U.S. industrial data for the 1948-1955 period. The author includes the effect of missed sales expectations in the model, and independently estimates manufacturing (raw material and work in process) and finished product inventories. He concludes that the extended period of inventory deficiency in manufacturing industries shown in the data supports the hypothesis that manufacturers use fraction inventory adjustments, and that, additionally, manufacturers tend to underestimate actual increases in sales. Hay (1970a) however, questions whether it is proper to assume that there exist costs associated with changing inventory levels other than those directly related to the costs of the necessary change in production.[12] The author argues that if it is the production costs that affect the change of inventory, then the proper model must generate fractional adjustments of the desired production, rather than of the desired inventory level. Thus, he proposes the following change in the model specification:

$$\hat{I}_t = a + bS_t, \tag{2.13}$$
$$\hat{Q}_t = S_t + \hat{I}_t - I_{t-1}, \tag{2.14}$$

where $\hat{Q}_t$ is the desired production at time t, which is then smoothed through:

$$\Delta Q_t = g(\hat{Q}_t - Q_{t-1}). \tag{2.15}$$

with $0 < g < 1$. Lovell (1971) retorts that without knowing the explicit cost functions used, the *a priori* rejection of the his model by Hay is unwarranted. Accepting production smoothing as a valid hypothesis, and citing that the recent (at the time) empirical literature examining the matter often concluded that a mix between production and inventory smoothing best fit the data. Hay, however, having presented this production-smoothing variant to the model, argues that even this specification is an oversimplification of reality and that a "proper" model must take into account more variables present in the cost calculation, as well as the interdependence between the variables. With this in mind, he develops a comprehensive empirical model where pricing, production, and inventory are estimated simultaneously (Hay, 1970b). This model is based on Holt et al.'s (1960) quadratic cost specification. In fact, it is one of the earliest macroeconomic studies based on this model (building upon early work by Belsley (1969) and Childs and Johnston (1967)), which –as mentioned above– is to become the "empirical workhorse" for macro empirical research.

The production smoothing model developed in Holt, Modigliani, Muth, and

[12] He, however, concedes that the flexible accelerator may be a useful concept in the case of investments in fixed capital, where the changes in inventory holdings do incur tangible cost.

Simon (1960) is particularly important because it introduces the concept of quadratic cost components. The intuition behind quadratic costs is very simple: given an optimal value of any of the variables, positive –as well as negative– deviations increase the costs. Furthermore, construction of such a cost function guarantees that an optimal solution exists and is a linear combination of the parameters. The original model is aimed at a plant-level decision maker and thus is defined in terms of regular payroll costs (C_P), hiring and layoff costs (C_H), overtime costs (C_O), and inventory costs (C_I). The total costs are then minimized for a finite horizon (T):

$$\min \sum_{t=1}^{T} \left(C_P + C_H + C_O + C_I \right), \tag{2.16}$$

where the inventory costs are modeled as

$$C_I = a_1 (I_t - (a_2 + a_3 S_t))^2, \tag{2.17}$$

where the *a's* are constants—note that this term includes, in essence, the "inventory accelerator" from the earlier models.

For the canonical basic production smoothing model as applied to aggregate economies, we turn to Blinder's (1986) paper, provocatively titled *"Can the production smoothing model of inventory behavior be saved?"*. In this paper, the authors briefly recount the "long and venerable" history of the production smoothing model, describe the mathematical structure of the production smoothing framework, and put forward the reasons why the model "is in trouble". In contrast with the original (Holt et al., 1960) model specification, macroeconomic applications of the production smoothing model attempt to maximize revenues, rather than to minimize costs. Let $R(x_t)$ be the revenue, dependent on the demand function x_t, and let $C(y_t)$ and $C(N_t)$ be the costs related to production (y_t) and holding inventory (N_t). Then, the firm wishes to maximize

$$E_0 \sum_{t=0}^{\infty} D^t \left(R(x_t) - C(y_t) - C(N_t) \right), \tag{2.18}$$

where the demand curve is assumed to be linear with stochastic shocks:

$$p_t = -\delta x_t + \epsilon_t, \tag{2.19}$$

where p_t is the price and ϵ_t a shock component[13].This results in a quadratic

[13]Unlike models in OR/OM, economic equilibrium models assume that demand is linear in price.

revenue function:

$$R(x_t) = p_t x_t = -\delta x_t^2 + x_t \epsilon_t. \tag{2.20}$$

We refer the reader to Blinder (1983), and Ramey and West (1999) for insights on the solution methods of such equilibrium problems. The production (y_t) and inventory (N_t) costs are assumed to be quadratic and defined thus:

$$C(y_t) = c_1 y_t + (1/2c) y_t^2, \tag{2.21}$$

$$C(N_t) = b_1 N_t + (b/2) N_t^2, \tag{2.22}$$

where c_1 and b_1 are constants, and c and b are the curvature parameters and are critical to the production smoothing issue. A large value of c represents a steep marginal cost curve for production, and thus a strong incentive for smoothing; a large value of b, on the other hand, represents large costs for changes of inventory and thus presents a strong disincentive for smoothing. Assuming i.i.d demand shocks, and that firms attempt to maximize the expected discounted value of profits, the author derives the analytical solution for the model and shows that the resulting variance of production is higher than the variance of demand, and that inventories and sales are counter-cyclical—exactly opposite to what the empirical data suggests. The authors offer two plausible amendments to the model that put the predictions in line with the observations: cost shocks, and serially correlated demand shocks. They show that if Equation (2.21) is modified in such a way that an additive i.i.d stochastic shock (Γ_t) is added to the production cost:

$$C(y_t) = (c_1 + \Gamma_t) y_t + (1/2c) y_t^2, \tag{2.23}$$

then production is more variable than sales, and the inventory and sales are pro-cyclical. A similar result is obtained when the demand shocks ϵ_t are highly serially correlated. However, the fact that these modifications to the model are able to replicate the empirical observations (that inventory and sales are pro-cyclical, and that production variance tends to be larger than sales variance) are not enough for this to be a convincing explanation. Empirically, even though cost shocks have been reported to affect inventory investment in the form of raw material price changes Blinder (1986), and as technology shocks Eichenbaum (1990), the authors concede that they are not a plausible complete explanation—these modifications impose strong assumptions on the model.

This implies that firms can set the sales price so as to achieve equilibrium between supply and demand.

Following the publication of Blinder's (1986), a large amount of research has been dedicated to understanding the puzzle, and to reconcile the production smoothing model with the empirical evidence. Ramey and West (1999) contributes to the discussion with a paper that documents the empirical puzzle, reviews the relevant literature, and proposes several ways to align the predictions of the production smoothing model with the empirical data. In particular, the authors propose two plausible explanations: the addition of persistent shocks to the cost of production, and the inclusion strong "accelerator" motive in the cost of carrying inventories. The authors define the total costs as:

$$C_t = .5a_0\Delta Q_t^2 + .5a_1 Q_t^2 + .5a_2(I_{t-1} - a_3 S_t)^2 + U_{ct}Q_t, \tag{2.24}$$

where the $a's$ are constants; Q_t is the period's production; ΔQ_t is the change in the period's production; I_t, end of period inventories; S_t, sales; and U_{ct}, a permanent cost shock. The first two terms in the model capture the costs of production and of changing production. The third term explicitly embodies the inventory holding and backlog costs. It is straightforward to see that this term is derived from the original production smoothing model of Holt, Modigliani, Muth, and Simon (1960) (Equation (2.17)). Like Blinder (1986), the authors assume a linear demand curve dependent on price, and solve the profit maximization problem to derive the linear decision rule. They characterize the response of the system by studying the influence of the different constants ($a_0 - a_3$) and the auto correlation of the demand and cost shocks. They find that the desired response of the system is achieved in the presence of demand shocks when there are large costs of adjusting production (a_2 small relative to either a_0 or a_1) and the inventory accelerator motive is strong (a_2a_3 large relative to a_0 and a_1), or when the marginal production costs are declining ($a_1 < 0$). Additionally, like Blinder (1986), they find that the required response of the system is also achieved in the face of persistent cost shocks of an auto-regressive nature. When they review the empirical evidence, however, they find that declining marginal production costs are not a convincing explanation, and that they do not have enough data to back one alternative over the other (or even over another alternative explanation). The paper ends with a call for the usage of different data to be able to test these explanations empirically.

Kahn (1986, 1992) provides a theoretical basis for including the inventory accelerator motive into the models in the form of stock-out avoidance and shows that under linear costs and i.i.d, or serially correlated, demand then the variance of production is higher than the variance of sales. Wen (2005), however, points out that from a macroeconomic perspective, Kahn's models make several restrictive assumptions that limit its applicability, specifically: the theory is based on a partial equilibrium model with exogenous demand and labor supply, as well as specific structural assumptions such as constant marginal production costs and no capital investment. Wen builds up on Kahn's work and develops a model based on general equilibrium theory that generalizes the latter's results. The author shows that Kahn's insights continue to hold, provided several conditions are met: either the marginal cost of production are constant and the inventory holding costs strictly positive, or costs of production are convex but inventory holding costs sufficiently large, or there exists an asset that dominates inventory investment in the long term expectation of returns so that there is no incentive to plan on holding inventories in the long run.

More recently, Ramey and Vine (2006) conduct a plant-level empirical study of the U.S. automotive manufacturing industry. The objective of their research is to understand why the variability of GDP decreased during the 90's and 00's as compared to the 80's. Using a variation of Holt et al.'s (1960) they find that a change in the scheduling strategy of the workforce is correlated with the drop of production variance at the plant level. Their findings suggest that plants have shifted their scheduling and the use of overtime hours has become more prevalent. They argue that this implies that the marginal production costs begun to be dominated by a more "traditional convex" function. From this it follows that non-convex lumpy margins, such as shift changes, present in the industry can be a cause of the variance of production being higher than that of sales—implying that the [inaccurate] assumption of convex production costs may play a role in the empirical problems of the production smoothing model.

2.2 *Analyzing the bullwhip*

In their analytical work, Lee et al. (1997a) identify four operational causes of the Bullwhip Effect. Under the assumptions of an Order Up To policy and a serially correlated demand process (AR(1)[14]), the authors analyze a single-stage, single-item, multi-period inventory model. The first of the operational causes of the Bullwhip Effect, demand signal processing, appears when past demand realizations are used to update the forecast of demand. When this is the case, the variance of orders is strictly larger than the variance of demand for positively correlated demands, and that this variance is strictly increasing in the replenishment lead times. Note that the authors do not make any assumptions on the forecast—they only require it be updated with past demand

[14] $D_t = d + \rho D_{t-1} + u_t$, where D_t is the demand in period t, $-1 \leq \rho \leq 1$, and u_t is an i.i.d error term with mean 0 and variance σ^2 .

information. To prove the influence of the next operational cause, shortage gaming, they develop an extended newsvendor model with one manufacturer and multiple retailers and show that if the combined retailer demand exceeds the manufacturer's production, and the latter allocates shipments proportional to demand, then the optimal order quantity for the retailer exceeds the optimal order quantity with no supply constraint. This implies that in a multi-stage supply chain, the rational behavior of managers –in the presence of *expectation* of supply shortage– induces the amplification of order variance. The intuition behind the third operational cause, order batching, is as follows: consider a supply chain where each player uses a periodic review replenishment system with full backlogging and the retailer(s) face stationary demand. Under this system, each retailer will order, every review period, the demand faced during the previous review cycle. If a firm supplies to more than one retailer, the orders from all retailers will arrive exactly at the same time, completely balanced (with no 2 orders arriving at the same time), or somewhere in the middle. Lee et al. (1997a) prove that the variability observed by suppliers is larger than the variability of retailer's orders. The final operational cause reported in the paper is the fluctuation of prices. The authors show that, when purchasing costs fluctuate between a high level (C^H) and a low level (C^L) with $C^L < C^H$, then the optimal inventory policy has the following form: When price is C^L steer the stock level towards S^L, and when price is C^H steer the stock level towards S^H, where $S^L < S^H$. This inventory policy results in $\text{Var}[O_t] > \text{Var}[D_t]$.

2.2.1 *The influence of inventory policies and demand forecasts*

Graves (1999) extends the demand signaling processing result shown above to the case where the demand is a non-stationary process defined by an integrated moving average (IMA) of order (0,1,1)[15]. For this demand stream, Graves derives an adaptive base-stock policy of the form

$$O_t = D_t + L(F_{t+1} - F_t), \tag{2.25}$$

where D_t is observed before calculating O_t, F_t is an exponentially smoothed forecast of demand for time t, calculated in period $t-1$, and L, the replenishment lead time. The author admits that this is not an optimal policy, but a rather reasonable adaptation of a base-stock policy to the non-stationary demand. In his analysis, Graves proves that under this setting, O_t is the same type of process as D_t, and that $\text{Var}[\text{O}_\text{t}] > \text{Var}[\text{D}_\text{t}]$; i.e., demand signaling processing also

[15] Also knwon as an ARIMA demand. It is defined by $D_1 = \mu + \epsilon_1$, and $D_t = D_{t-1} - (1-\alpha)\epsilon_{t-1} + \epsilon_t$ for $t > 1$; μ and α are known parameters, and ϵ_t is a series of i.i.d normally distributed random variables, with $\text{E}[\epsilon_t] = 0$ and $\text{Var}[\epsilon_t] = \sigma^2$.

generates bullwhip of orders under this setting. Perhaps most interesting is his proof that, in multiechelon contexts operating under this adaptive base-stock policy, there is no advantage in sharing downstream information with upstream players because the optimal forecast is an exponentially smoothed time series that depends on the orders observed upstream, and is not improved by knowing the actual realization of downstream demand—provided the upstream player knows the parameters of the customer demand process.

Hoberg et al. (2007b,a) use control theory to explicitly investigate the influence of inventory policies and the exponential smoothing parameter, α on demand and inventory amplification. Both papers analyze a two-echelon supply chain and compare the installation-stock and echelon-stock versions of an order up to policy with full adjustments, inventory coverage C, and exponentially smoothed forecasts—effectively the policy described in Equation (2.5), where $\gamma_I = \gamma_P = 1$. The installation-stock version of the OUT policy takes into account the inventory level at each individual echelon to generate new orders; the echelon-stock version of OUT the policy takes aggregated information into account: The upstream echelon uses downstream order and demand information to inform its decision-making process. In Hoberg et al. (2007a) the authors concentrate on the stationary performance of the system (measured by the bullwhip ratios, BW_O and BW_I) and in Hoberg et al. (2007b), they concentrate on the transient performance of the system (measured by the ITAE of orders and inventories). Additionally, Hoberg et al. (2007a) also analyze an inventory-on-hand policy, where inventories alone are taken into account for order generation—equivalent to Equation (2.5) with $\gamma_I = 1$ and $\gamma_P = 0$. This policy is found to be unstable for positive lead times. In accordance with traditional multi-echelon research (Clark and Scarf, 1960) the authors find that the echelon-stock policies offer superior performance for virtually every possible parameter combination and performance measure.[16] Analyzing the influence of the exponential smoothing parameter, the authors find that the stationary performance (measured through BW_O and BW_I) is decreasing in α, but the transient performance (ITAE_O and ITAE_I) is increasing in α. Thus, the usage of exponentially smoothed forecasts imposes a trade-off in the system design; the authors recommend low values of α when demand is stationary, to avoid amplification, and high values of α when demand is not stationary, to adapt to changes in the demand's mean.

Dejonckheere et al. (2003) study the influence of different forecasting mechanisms on the performance of the OUT policy with full adjustments (Equation (2.5), with $\gamma_I = \gamma_P = 1$), and the performance of a subset of the OUT

[16] The exception being the response to certain specific sinusoidal demand patterns (purely seasonal demands) where installation-stock policies perform marginally better.

with fractional, but equal, adjustments of inventory and pipeline feedback loops (Equation (2.5), with $0 \leq \gamma_I = \gamma_P \leq 1$). As mentioned in §2.1.2 this family of policies is named DE-APVIOBPCS in honor of Deziel and Eilon, whose work was the first to suggest fractional adjustments as a way to reduce amplification. The OUT policy with full adjustments is analyzed with exponentially smoothed forecasts, moving average forecasts, and with demand signal processing (Lee et al., 1997a) as a forecast. The authors prove that OUT policies, no matter what forecasting mechanism is used, will always result in the Bullwhip Effect. They suggest the adoption of the DE-APVIOBPCS policy in cases where the variability of orders results in excessive costs, for they allow the decision maker to eliminate variance amplification at the cost of responsiveness.

Jakšič and Rusjan (2008) perform a single-echelon transfer function analysis of the performance of a family of replenishment policies derived from the work of Bowman (1963). The general form of the family of policies is:

$$O_t = \hat{D}_t + (1-\gamma)(O_{t-1} - \hat{D}_t) + \beta(IP_t^T - IP_t), \tag{2.26}$$

where O_t is the order placed at time t, $\hat{D}_t$ the forecast of demand at time t, IP_t the inventory position at time t, and IP_t^T the desired inventory position at time t, calculated through $IP_t^T = \hat{D}_t T_L + k\hat{D}_t\sqrt{1+T_L}$ with k a safety factor and T_L the replenishment lead time. The choice of the arbitrary smoothing factors γ and β define the replenishment policy: when $\beta = 0$ and $\gamma = 1$, then the policy reduces to ordering the forecasted demand every period; when $\beta = 0$ and $0 \leq \gamma \leq 1$, the policy incorporates smoothed information of past orders while ignoring the inventory position; the opposite is true when $0 \leq \beta \leq 1$ and $\gamma = 1$; the policy reduces to a conventional base stock policy when $\beta = \gamma = 1$; and the policy at its most general is defined by $0 \leq \beta \leq 1$ and $0 \leq \gamma \leq 1$. The authors perform numerical experiments using different demand streams and, confirming prior findings, show that the inclusion of the inventory position in the ordering rule introduces variance amplification at specific frequencies, depending on the system's parameters. They perform a cost analysis by introducing fixed ordering costs and variable holding and shortage costs and conclude that, under certain conditions, ignoring the inventory position can bring about overall savings. The authors do not report, however, the cost parameters used for their study.

2.2.2 *The influence of behavior*

Sterman (1989) argues that the poor performance of individuals playing the beer game arises from "misperceptions of feedback"; that is, humans are inherently unable to incorporate the feedback structure of a dynamically complex system into their decision making—specifically, we seem unable to completely account for the cumulative pipeline when placing orders. In other words, he argues that the Bullwhip Effect is essentially a behavioral problem. His experimental work, however, being carried out in the lab, is unable to provide any certainty about this effect being present in higher levels of aggregation. Sterman shows that individual human biases *can* generate the Bullwhip Effect, but he cannot show that the Bullwhip Effect in the real world *is* caused by human biases. His work, nevertheless, inspired valuable research on the behavioral causes of the Bullwhip Effect. In a series of related studies, Croson and Donohue (2005, 2006) use a controlled version of the beer distribution game where they suppress the known operational causes of the Bullwhip Effect: Order batching, shortage gaming, forward buying, and errors in demand signaling. The authors test hypotheses related to information sharing and find (a) that the Bullwhip Effect and the under-estimation of the pipeline persist when the end market demand is stationary and known by all players; (b) that the Bullwhip Effect and the under-estimation of the pipeline persist, but are reduced, when dynamic inventory information is relayed to all players (i.e., every player has complete knowledge of the inventory level of every other player) and that the magnitude of the performance improvement is larger for upstream players than for downstream players (Croson and Donohue, 2006); and (c) that sharing *only* downstream inventory information reduces the oscillations in the supply chain –specially upstream oscillations– whereas sharing *only* upstream inventory information does not result in the reduction of oscillations at all (Croson and Donohue, 2005).

Diehl and Sterman (1995) attempt to disentangle the effects of feedback processes and time delays in human decision making through an experimental study that combines econometric fitting of decision-making models with subjective data taken from "notebooks" given to the subjects, where they annotate calculations and other general comments on their experience. The experimental set up is based on a single-echelon system to isolate the intra-echelon dynamics and incorporates time delays in the production process, and an explicit feedback process in the demand generation.[17] The performance of the subjects decreased as time delays and feedback gain increased. In the more difficult conditions (long delays and high feedback gain), human players performed worse than a naive, no-control rule. The authors conclude that the poor performance of human subjects stems from 2 different mechanisms: On the one hand, people over-simplify complex tasks when creating mental models, often excluding side effects, feedback structures, and other dynamic effects,

[17] Demand is a combination of an exogenous process –a random walk– and an endogenous process –a proportion of the prior period's production fed back into the demand process–.

resulting in an inherently human difficulty for understanding accumulation;[18] on the other hand, we simply cannot solve complex systems of non-linear differential equations intuitively. The former can be improved by training (see Sterman (2010) for a study on the benefits of formal training in the human understanding of accumulation); the latter is a human limitation.

Croson et al. (2014) look at the effect of human behavior in the bullwhip in a different way. They study the influence of "coordination risk" and "coordination stock" in the inter-echelon dynamics of a modified beer distribution game set up. The thinking behind these concepts is that there is a risk involved when collective performance (the bullwhip of the entire supply chain) depends on individual decisions (orders at each echelon) that are not known with certainty by all the players—i.e., coordination risk; and that holding extra inventory as a buffer against this risk can mitigate the oscillations generated by the Bullwhip Effect—i.e., coordination stock. The authors test a series of hypothesis related to these concepts by conducting different experiments with undergraduate and graduate students of diverse disciplines. They find that the Bullwhip Effect and pipeline under-weighting persist in all the experimental variants, thus providing evidence that the Bullwhip Effect is –to an extent– a behavioral problem. Additionally, they identify "coordination risk" as another mechanism through which the bullwhip manifests itself, and propose two mechanisms that, reducing this risk, mitigate the bullwhip. These mitigating mechanisms are (1) sharing knowledge about optimal decision rules among all players and (2) keeping "coordination stock", essentially a physical buffer against oscillations. Even though these strategies reduce the amplitude of oscillations, they appear not to reduce the behavioral trait of pipeline under-estimation.

2.2.3 *Rogue seasonality*

Rogue seasonality is, of all of Forrester's (1958) contributions to the theory, the least explored manifestation of the Bullwhip Effect. It is easy to see why this aspect is comparatively under-explored in the literature: By definition, rogue seasonality is a periodic phenomenon; it cannot be identified through stationary measures, and its analysis in the time domain is qualitative at best—the overwhelming majority of literature related to this phenomenon concerns frequency-domain analysis. If we analyze the replenishment policy of a single echelon in a supply chain in the frequency domain, for example through control theory methodologies, the intuition behind the appearance of rogue seasonality (also called fake business cycle by Forrester) is easy to understand: every system

[18] Rather than understanding stocks increase when inflow exceeds outflow, people tend to use a *correlation heuristic*, concluding that a system's inputs and outputs are correlated (Sterman, 2010).

has a "spectral signature" that depends on the replenishment policy, the type of forecasting method, and the specific parameters; this "spectral signature" is immediately seen in frequency response plots; certain frequencies are amplified more than others. It follows then, that if the input is uniform[19], then the output will adopt the spectral signature of the system. We are not aware of any study that analyses the seasonality of a system's output in the time domain; as we have recounted in this review, much of the research on the Bullwhip Effect has focused on the amplification of order variance—even research focusing on the frequency domain, such as control theoretic models. Kim and Springer (2008) perform a mathematical analysis of Sterman's (1989, 2000) system dynamics single-echelon supply chain model with the aim of defining specific conditions under which rogue seasonality appears. The authors propose a test for the endogenous seasonality generation: If, after being affected by a demand shock, the system's inventory and pipeline oscillate before converging to an equilibrium value, then the system generates endogenous oscillations. The authors derive a series of conditions under which endogenous seasonality is avoided; in general, decreasing the replenishment lead time, decreasing the pipeline adjustment time, or increasing the stock adjustment time, increases the chances of rogue seasonality avoidance. This result, however, depends on a crucial assumption of the system dynamics model specification: lead times are approximated by first order exponential delays. Mathematically, this results in a second order system for all lead times; in second order systems, we can find whether the response is oscillatory by finding whether the roots of a second degree polynomial are real or complex conjugates—the latter introduce oscillations. If the first order lead time assumption is dropped, however, the system's order depends on the lead time, and therefore an analysis as the one presented in Kim and Springer (2008) cannot be conducted for general lead times (see Chapter 4 for more details on the analysis of high-order systems).

We can, nevertheless, analyze endogenous seasonality in systems with realistic lead time assumptions using frequency domain analysis. Thornhill and Naim (2006) illustrate an application of spectral principal component analysis (SPCA) to the detection of rogue seasonality in a supply chain that can be applied to empirical, as well as simulation, data. The approach they adopt is a direct application of the methods Thornhill et al. (2002) develop for the detection of oscillations in the processes of a chemical refinery, and is based upon clustering the data according to the similarity in frequency responses. The method comprises of several steps: First, we mean-center each of the time series data we want to compare, and calculate its power spectra (i.e., its frequency-domain

[19] As is the case of a normal demand, where by definition all frequencies are equally represented.

representation) using Fast Fourier Transforms (FFT); we then normalize the power spectra to allow for comparisons of different systems; finally, we use principal component analysis (PCA) on the data to reduce the dimensionality of the frequency response to 2 or 3 principal components, which can be plotted and clustered. For more details regarding this methodology, we refer the reader to Chapter 5.

2.3 *Taming the Bullwhip*

Lee et al.'s (1997a) paper, in addition to showing the operational causes of the Bullwhip Effect, identified possible strategies for its mitigation such as information sharing , vendor managed inventory (VMI), and the reduction of replenishment lead times. In this section, we describe work carried out since with the objective of mitigating the Bullwhip Effect, or –in Lee's own words– the *taming of the bullwhip*. The section is split in two: In §2.3.1 we survey the literature that seeks bullwhip reductions through the optimization of the parameters of individual systems, and in §2.3.2 we survey the literature concerned with optimizing the supply chain performance through the sharing of information across firms.

2.3.1 *Optimizing intra-echelon performance*

In the control theory world, the AP(V)IOBPCS framework (to describe a generalized OUT policy) is the most widespread tool for the analysis of inventory systems. Among this framework, it is well known that the subset of DE-AP(V)IOBPCS policies have desirable characteristics. Such systems are always stable, and generate a response with only limited oscillations (see Dejonckheere et al., 2003 and Disney et al., 2006a). In Disney et al. (2006a), the authors use a DE-APVIOBPCS policy to investigate the impact of parameter combinations in bullwhip reduction when the demand is a generalized ARMA process. They argue that setting the parameters of the inventory policy needs to be done in a case by case basis, since the desired response depends on the "peculiarities" of each supply chain; the reduction of variability is not always the ultimate goal—inventory and customer service levels are important dimensions in the design of a system. Illustrating this point, Disney et al. (2013) presents a case study based on the implementation of a decision support system based upon the APVIOPCS policy for the control of Lexmark's printer toner's production line. They report that after implementing the DSS the bullwhip for orders was reduced five-fold and the bullwhip for inventory was reduced tenfold, without a hit in customer service level. They do not, however, report on the changes in the inventory level itself.

Not all control-theoretic studies of inventory-production systems, however, use the AP(V)IOBPCS framework. In a study that predates it, Bertrand (1980) analyses the variance amplification of inventories and order releases in the diffusion department of a semiconductor plant. Due to the cyclic nature of the manufacturing process, the production is modeled as a multi-stage system with serially connected processes and intermediate inventories. Uniquely, stochasticity in this system is generated by random yields of the intermediate processes rather than by random demands. Just as Disney et al. (2013), the motivation for this study is entirely practical: Several months after implementing a control rule based upon the line-of-balance technique, personnel started complaining. The variance of both inventories and order releases was such that at times the available capacity was too low for the requirements, while at other times the requirements were very small. This behavior was due to the adoption of a control rule that was optimized for steady state—its dynamic behavior was not analyzed before the implementation phase. Bertrand (1980) uses the z-transform technique to derive expressions for the steady state behavior and the impulse response of the system; studies the trade-off between steady-state and dynamic behavior, and proposes a new decision-rule that incorporates smoothing of the target deviations. He concludes by advocating thorough analyses of the dynamic behavior of control rules, citing control theory as a solid methodology for the purpose.

Li et al. (2013) turn their focus to another widespread tool in the analysis of inventory systems: the use of exponentially smoothed forecasts. In this study, the authors prove that simple exponentially smoothed forecasts, as well as forecasts that follow Holts method, always result in bullwhip when using an OUT policy. They approach the question of bullwhip reduction by adopting a dampened trend forecast (both Holts and simple exponentially smoothed forecasts are a generalization of a dampened trend forecast). They show that using this forecasting technique, the appearance of the bullwhip depends on the demand pattern, and can be avoided under certain conditions. Their analysis generates sufficient conditions for bullwhip generation and necessary conditions for bullwhip avoidance. They call for more attention to be given to dampened trend forecasts as a way of reducing the bullwhip. This view, contrary to the use of smoothing policies, relies on obtaining "better information" to be used as an input to the policy rather than adopting a robust policy.

2.3.2 *Information sharing in supply chains*

It is no surprise that information sharing is often proposed as a remedy for the Bullwhip Effect—the latter is, after all, caused in part by the transmission of distorted information up a supply chain (Bray and Mendelson, 2012). Information sharing has received substantial attention in recent years, in part because advances in technology have enabled the practical application of such systems in the industry. What was implausible at the moment of Forrester's first unveiling of the bullwhip (e.g. the sharing of real-time inventory information across remote locations) is nowadays attainable.

The information sharing literature can broadly be classified (1) according to the source of the information that is to be shared, and (2) according to the relationship among the various information-sharing players.

Regarding the source of information, we can identify a stream of literature that analyses the potential benefits of sharing *downstream information* with upstream players, and conversely a stream that analyses the potential benefits of sharing *upstream information* with downstream players. Regarding the relationship among the various supply chain players, a large body of literature assumes the point of view of a central planner of sorts, with the objective of optimizing the supply chain as a whole. This stream extends the OR/OM literature using concepts, models, and assumptions from multi-echelon inventory theory; the analysis is often centered in quantifying the benefits of information sharing with regards to supply chain costs and/or service performance metrics.

The complementary stream, on the other hand, assumes that a supply chain comprises independent firms with private information. Under this paradigm, incentive issues arise: Due to asymmetries in the quality of information that is to be shared (and the resulting benefits), an entire new set of research questions crops up. In this view, it is not immediately clear whether different firms have sufficient incentives to share information. Thus, the analysis shifts towards deciding whether information sharing is an equilibrium solution of a non-cooperative game; it is often centered in deriving the conditions under which every independent firm in the supply chain benefits from sharing the information they posses.

We refer the reader to Chen (2003) for a comprehensive review on information sharing in supply chains according to the aforementioned taxonomy. In the remainder of this section, we pay particular attention to the literature that concerns the sharing of *downstream information* with upstream firms with the Bullwhip Effect as a performance measure.

Chen et al. (2000) consider a two-stage supply chain model with a serially correlated, AR(1), downstream demand stream (see footnote 14). Each stage in the supply chain uses an OUT policy and a simple moving average forecast to update, every period, its estimates of the demand mean and variability after. They show that under these settings, sharing complete information of the downstream demand will reduce the bullwhip observed upstream, but will not completely eliminate it.

Lee et al. (2000) consider a similar setting, with AR(1) demand, to study the benefits of sharing downstream demand information in a two echelon supply chain. In this study, however, the authors assume that both the downstream firm's demand process and its ordering policy are fully known by the upstream player. This allows the upstream to update the estimates of future orders using the previous period's orders and their knowledge of the underlying demand structure (as opposed to a simple moving average forecast). In this study, the effect of information sharing on the Bullwhip Effect is not explicitly measured; instead, they quantify the effect on the upstream order variability[20]. The authors show that sharing downstream demand information reduces the variability of upstream orders and is thus beneficial. They also note that the longer the lead times and the more variable, or more correlated, the demand is, the larger the potential benefits of information sharing are.

Raghunathan (2001), however, argues that the benefits of information sharing found by Lee et al. (2000) become insignificant if the manufacturer makes "intelligent use of available information". He shows that the manufacturer can get a highly accurate estimate of future demand using the entire order history, rather than just the last period's order, to update the forecasts. Under this setting, the benefits of information sharing are negligible because the manufacturer already has full information about the demand process, and the order history contains information about the demand observed by the retailer.

Gaur et al. (2005) extend and generalize these results to the more general ARMA (p,q) demand process[21]. They generalize the results of Raghunathan (2001) by presenting a rule to determine whether the order information can be used to infer the demand information. This allows the manufacturer to determine under which situations there is additional value in information sharing.

[20] Assuming that downstream orders are not affected by information sharing, this measure is indeed a proxy for the Bullwhip Effect.

[21] An ARMA (p,q) demand process follows $D_t = \mu + \rho_1 D_{t-1} + \rho_2 D_{t-2} + \cdots + \rho_p D_{t-p} + \epsilon_t - \lambda_1 \epsilon_{t-1} - \lambda_2 \epsilon_{t-2} - \cdots - \lambda_q \epsilon_{t-q}$ where ϵ_t are uncorrelated random variables of mean zero and variance σ^2, and $\rho_1, \ldots, \rho_p$ and $\lambda_1, \ldots, \lambda_q$ are known constants.

Chen and Lee (2009) argue that the assumption that the manufacturer has full knowledge of the retailer's demand process and ordering policy is, in reality, an aggressive assumption. They note that if these assumptions are relaxed, then the value of point of sales data (retailer demand) may be of limited use. Chen and Lee (2009) suggest that, in such a case, the sharing of the retailer's forecast may be of greater value. In this study, the authors adopt a most general model of demand evolution, the Martingale Model of Forecast Evolution (MMFE[22]) and a general order up to policy structure. The authors derive optimal ordering policies that minimize the total supply chain costs when the retailer's forecast information is shared; they argue that by sharing forecast information, the retailer is supplying the manufacturer with all the information it needs (thus, dropping the aggressive assumption that manufacturers have full knowledge about the demand and ordering processes of the retailer). Additionally, they note that by receiving the downstream forecast information, the manufacturer is able to separate between order variability and order uncertainty. In doing so, order variability is not the key cost driver anymore, instead, it is the uncertainty of the order revisions that drives the manufacturer's cost. This, they contend, calls for a rethinking of the use of order variability as the common quantification of the Bullwhip Effect—the empirical work of Bray and Mendelson (2013), discussed in §2.1.5 is an example of this approach.

From a control theoretic perspective, Dejonckheere et al. (2004) study the value of information sharing in traditional and generalized OUT policies. They investigate the influence of the forecasting mechanism and the assumptions of demand distribution (they derive expressions for the bullwhip using i.i.d normally distributed demands, and show the response to general demand distributions through frequency plots). Their analysis confirms the finding that information sharing is capable of reducing –but not completely eliminating– the bullwhip depending on the supply chain's parameters. They further quantify the reduction in bullwhip (under i.i.d normal demands) as one travels upstream in a supply chain as going from a geometric increase in the bullwhip ratio when no information is shared, to a linear increase when upstream players know the customer demand data. Furthermore, they find that in the case of the smoothed OUT policy, the variability of the bullwhip ratio (comparing the ratios at different echelons) is also reduced.

Cannella and Ciancimino (2008) analyze the impact of the parameters of a DE-APVIOBPCS design under progressive information sharing strategies through

[22]The MMFE demand process follows $Dt = \mu + \sum_{i=0}^{\infty} \epsilon_{t-i,t}$ where $\epsilon_{t-i,t}$ is the incremental information obtained in period $t-i$ with respect to demand D_t. $\epsilon_{t-i,t}$ is mutually independent, stationary, and normally distributed with mean zero and variance σ^2.

a continuous time differential equation analysis. Considering a supply chain derived from the beer distribution game setting, the authors conduct a series of numerical experiments to test the performance of 3 different information sharing strategies (no information sharing, sharing of customer demand information, and vendor managed inventory). Their findings suggest that using the smoothing parameters of the replenishment policy by themselves to reduce the bullwhip can have the undesirable effect of worsening the service level; that information sharing helps in reducing the bullwhip; that the deeper the information sharing strategy is, the larger the bullwhip reduction; anf, finally, that –in general– the bullwhip reduction due to information sharing are larger than those achieved by the smoothing of orders.

To illustrate the practical application of information sharing, Kelepouris et al. (2008) report on a computational study performed with data obtained from a Greek retail grocery store. They consider a two echelon supply chain (warehouse and retailer) using an OUT replenishment policy with exponentially smoothed demand forecasts and perform simulations to understand the influence of lead times and information sharing on the system's performance. Since their study is numerical in nature, they relax several assumptions that are common in the analytical literature (they do not allow either back-orders or returns in the model). They find that lead time has a significant impact in both the oscillations and fill rate of the system, and that sharing information of the retailer demand reduced the bullwhip of the system because it allowed the warehouse to generate more accurate forecasts. Additionally, the authors show that the sharing of information also lets the warehouse reduce the distortion stemming from order batching at the retailer.

From an implementation perspective, De Kok et al. (2005) present the development of an advanced planning and scheduling system that supports collaborative planning of operations between Philips Semiconductors and one of its customers, Philips Optical Storage. In this study, the reduction of the Bullwhip Effect is explicitly set forth as one of the main objectives behind the development of the collaborative planning project.

In contrast with the previously discussed analytical literature, the main focus of this study is the *implementation* of collaborative decision making that incorporates information sharing. The authors base their collaborative planning toolset on synchronized based stock policies (De Kok and Fransoo, 2003) that take into account stochasticity in demand, lead times, and yield through the concept of safety lead times.

To feed the planning tool, all partners collect and share live data on actual stocks, scheduled receipts, and material in transit. Thus, the outcome of the planning tool is a release plan for all the firms involved (Philips Optical Storage, Philips Semiconductors, and contract manufacturers). Previously, schedules were maintained individually per firm. Once the parameters are calculated, the generation of short-term feasible plans does not involve an optimization, rather just a calculation step that is essentially trivial. The authors stress that the biggest roadblock to adoption was not an implementation difficulty, rather to generate trust in the generated plans. The impact of the adoption of collaborative planning is estimated as a yearly savings of 2% of the total turnover, derived from inventory and obsolescence reductions. Key stakeholders assert that the reduction of information lead time, as well as increased visibility of up-to-date information enables decision making "based on facts".

There are 10^{11} stars in the galaxy. That used to be a huge number. But it's only a hundred billion. It's less than the national deficit! We used to call them astronomical numbers. Now we should call them economical numbers.

Richard Feynman

Chapter 3
De-stocking and the Bullwhip: an Empirical Investigation

The world economy experienced a sudden, severe and synchronized collapse in late 2008. The magnitude of the drop in global trade was the largest since World War II, it was the steepest in recorded history, and it was synchronized: all 104 nations where data is collected by the WTO experienced a drop in imports and exports during the second half of the year (Baldwin, 2009). Following the public collapse of the financial system (starting with the Lehman Brothers bankruptcy in September 2008), firms all over the world observed substantial demand disruptions; sales plummeted across the board, and panic spread. While many consumer markets remained relatively stable (exceptions being consumer durables and capital goods), the manufacturing sector observed almost instantaneous demand drops (Dooley et al., 2010).

In crises such as these, managers are pressured to improve the financial position of the company at the same time that demand levels are dropping dramatically. This typically leads to strategic decisions such as reducing inventories (to reduce the level of working capital), downsizing (to reduce operational expenses), and closing manufacturing facilities (to reduce fixed assets). These decisions, however, have substantial operational consequences when demand increases at a later stage: the reduction of inventory levels, workforce, and manufacturing facilities are decisions that require significant time to be reversed. If the situation that triggered such decisions is temporary and demand recovers faster than the speed at which firms can react, lost sales and general problems with inventory management will appear. Knowledge about the underlying dynamics behind the demand slump is therefore needed to avoid costly mistakes.

3.1 Introduction

In this chapter, we argue that firms reacted to the 2008 financial crisis by reducing their working capital targets and –because it was global and synchronized– this reaction introduced a significant shock in the world's supply chains, essentially creating an inventory-driven Bullwhip Effect. To test our hypothesis, we adopt supply chain modeling, experimentation, and validation methods based on theory from the experimental work by Sterman (1989) and Croson and Donohue (2006), originally focused on the appearance of the Bullwhip Effect following demand shocks in a laboratory setting. We develop 4 different supply chain models for a major chemical company in the Netherlands and validate them with demand data from the crisis period. In terms of methodology, our work distinguishes itself from previous studies on inventory dynamics by using extensive empirical data, framing the Lehman Brothers collapse as a natural experiment. We specifically distinguish between the direct estimation of the operational model parameters such as lead times, and the econometric fitting of behavioral parameters such as stock adjustment times. In terms of theory, we model aggregates of companies at a particular level of the supply chain in a particular region rather than individual decision makers (as is common in experiments) or firms (as is common in much of the system dynamics literature in supply chain management). The crisis time-frame, through the resulting synchronization in managerial objectives, gives us the opportunity to link aggregate and individual human behaviors.

Our results show that demand drops in the respective end markets were not severe enough to explain –by themselves– the wild dynamics observed upstream. Moreover, we show that the combination of declining end-markets and the appearance of a synchronized inventory shock successfully account for a significant portion of the observed long and short term dynamics. In this view, exogenous end-markets drive the overall long-term evolution of sales, while endogenous behavior (such as the inventory decisions taken as a consequence of the crisis) primarily impacts the short term dynamics.

The contribution of this chapter is thus threefold: (1) We identify the 2008 financial crisis as a natural experiment that –with the introduction of a synchronized inventory shock– effectively controls for the masking effects of aggregation. This allows for the usage of a system dynamics framework based on the Bullwhip Effect literature whereupon we model aggregate echelons. (2) We introduce a de-stocking hypothesis capable of explaining the demand evolution observed by upstream companies following the bankruptcy of Lehman Brothers. (3) We identify the importance of both consumer end-markets, and ordering behavior in the evolution of demand patterns through time. By explicitly modeling separate structural, operational, and behavioral parameters, this study quantifies their contribution to the observed transient behavior and allows for a comparison with results obtained from experimental studies on individual human decision making.

The remainder of this chapter is organized as follows: In Section 2 we introduce the historical views of inventories as seen in the economics literature, we identify the challenges inherent in the study of inventories as part of aggregate models, and develop our de-stocking hypothesis. Section 3 introduces the methodology and model formulation, extending prior experimental work and framing it in the crisis time-period by explicitly modeling behavioral managerial decisions in the form of a reduction of inventory targets. In Section 4, the echelon model is used to model four different supply chains, we use empirical data to calibrate and validate the models, and develop forecasts for these supply chains. We then formulate an alternative model –sans the de-stocking hypothesis– to study the appropriateness of this hypothesis. We conclude in section 5 with a series of managerial insights.

3.2 *Background and Hypothesis Development*

When looking at the link between inventories and macro economic developments, Blinder and Maccini (1991) point out that interest in inventory behavior seems to follow cycles, not unlike the economy we attempt to explain. Indeed, we observe that research on the role of inventories in the economy peaks throughout history following extraordinary economic happenings such as the post-war period, the late seventies oil crisis, and –relevant to current developments– the financial crisis of 2008.

We refer the reader to Fitzgerald (1997) and Blinder and Maccini (1991) for comprehensive reviews of over 50 years of discussions on inventory theory in the economics discipline. In his work, Fitzgerald (1997) identifies inconsistencies between theory and data, and the subsequent attempts of researchers to eliminate these discrepancies from their models. Blinder and Maccini (1991) summarize the opposing views of micro and macro economists with regards to the role of inventories: the former discipline sees them as a stabilizing factor, whereas the latter sees them as a de-stabilizing one. Despite these fundamental disagreements, Feldstein and Auerbach (1976) point out, inventory fluctuations have long been recognized as a major endogenous force in American business cycles. In their experience, irrespective of the conceptual contradictions between contemporary models and the real-life processes behind them, most studies of inventory behavior note that about 75 percent of the cyclical downturn in gross national product (from peak to trough) can be accounted for by the reduction of business inventories. Recognizing these conceptual difficulties, Lovell (1994) reflects upon the inherent challenge of trying to reconcile these views. He poses a series of questions that –for all the body of research available– remain open to this day: "(...) Do firms actually attempt to smooth production? Is an empirical analysis of industry-level data enough? Is it necessary to analyze firm-level data in order to explain these effects?". These questions read as a research agenda on the mechanisms behind empirical observations on both macro and micro levels, recognizing, amongst other issues, the potential masking effects of aggregate data. In the operations management (OM) literature, inventory theory is often developed in a stylized manner; with strong assumptions that favor mathematical tractability over the inclusion of the myriad factors that are present in real life. The objective of these simplifications is to develop managerial insights that are both rigorous, and useful in the real world. In an exploratory study, Rumyantsev and Netessine (2007) find evidence that many insights from classical inventory models survive aggregation and do, in fact, hold up when analyzing empirical data.

The dynamics that stem from the interactions of subsequent echelons along a supply chain have been extensively studied in the OM literature. The fact that relatively small shocks can introduce severe instabilities in entire systems was shown by Forrester (1958), and is a central idea behind the Bullwhip Effect. The Bullwhip Effect has long been analytically and experimentally understood, and its effects and causes have sparked a great amount of research that has delivered valuable managerial insights (Sterman, 1989; Lee et al., 1997b; Croson and Donohue, 2006). However,even though the Bullwhip Effect itself is significant at the firm level (Metters, 1997; Fransoo and Wouters, 2000; Bray and Mendelson, 2012), attempts to empirically quantify the effect at higher aggregation levels have not been successful: studies have failed to prove it statistically significant at an industry level (Cachon et al., 2007; Bu et al., 2011). The lack of clear empirical evidence is attributed to the influence of factors present in government statistics such as their high level of aggregation (Chen and Lee, 2012), and seasonal adjustment (Gorman and Brannon, 2000). Furthermore, as Rumyantsev and Netessine (2007) point out, extending many structural properties from single-product, single-echelon models to higher aggregation levels also requires the assumption that products be homogeneous and their inventory control be synchronized.

With this in mind, the financial crisis of 2008 allows us to study empirical data in a different way. Following the bankruptcy of Lehman Brothers on September 2008, the financial world found itself in turmoil; credit dried up almost instantly and many companies in the world shifted their financial priorities according to the "cash is king" motto: liquidity became essential. Freeing up cash in the short term through inventory divestment is one strategy that can be followed by companies in times of distress (Sudarsanam and Lai, 2001). In a recent work, Pesch and Hoberg (2013) conduct an empirical study that shows that firms in financial distress reduce their inventories as part of their turnaround strategy: 70% of the firms in their sample reduce their inventories, with a median reduction of 9.4% of all inventories. We hypothesize that firms all over the world reacted to the financial collapse by significantly reducing their inventory targets. This, combined with the extraordinary synchronization observed during the the period (Alessandria et al., 2010) and the ever increasing influence of supply chain dynamics in the global economy (Escaith et al., 2010), introduced a synchronized, endogenous, inventory shock that generated an inventory-driven Bullwhip Effect. Early studies following the financial crisis seem to confirm this view in the manufacturing sector (Dooley et al., 2010). Using the collapse as a natural experiment, we model supply chains at an aggregate-echelon level, use exogenous end market data to drive those models, and validate them with primary empirical data collected at a major dutch chemical company.

3.3 *Theoretical Background and Model Structure*

In this section, we present our echelon model based upon Sterman's managerial decision making and supply chain models (Sterman, 1989, 2000) and follow with an introduction to the de-stocking logic we use to model the hypothesized reaction to the credit crisis.

3.3.1 *Echelon model*

An echelon model consists of three sectors (see Figure 3.1): the forecasting and orders sector tracks the incoming customer orders, maintains the echelon sales forecast, and generates material orders. The production sector regulates inventories and production, and the delivery sector keeps track of customer deliveries and backlogs. The model assumes no lost sales, and is based on continuous time system dynamics simulations. There is no sequence of events as such; cause and effect relationships are modeled by differential equations (i.e., we model rates of change), and products are modeled as continuous flows (demand is an outflow, incoming orders an inflow).

Because these echelon models are intrinsically linked to one another (deliveries from one echelon will become material receipts for the echelon immediately downstream in its supply chain), each of the parameters we define has a subscript $[n = (1, ..., N)]$ representing its place in the supply chain. We number echelons from downstream to upstream: the most downstream echelon being 1 and the most upstream N. In the case of diverging supply chains, where one echelon can potentially have several direct customers, we introduce a second number following a period, indicating the existence of other parallel echelons in the supply chain.

Table 3.1 *Definition of model parameters*

S_n	on-hand stock at echelon n
PL_n	Pipeline inventory at echelon n
W_n	work in process stock at echelon n
P_n	production rate of echelon n
F_n	sales forecast at echelon n
L_n	incoming delivery lead time at echelon n
PT_n	production time of echelon n
$\tau_n(PL)$	pipeline inventory adjustment time at echelon n
$\tau_n(S)$	stock adjustment time at echelon n
$\tau_n(F)$	forecast adjustment time at echelon n
$\tau_n(L)$	expected delivery delay at echelon n
$\tau_n(L)$	minimum time to fill orders at echelon n
$\hat{C}_n$	desired on-hand inventory coverage at echelon n
$\hat{D}_n$	desired delivery rate echelon n
$\hat{S}_n$	desired on-hand inventory at echelon n
$\hat{PL}_n$	desired supply ilne at echelon n
$O_n(PL)$	pipeline inventory adjustment of orders at echelon n
$O_n(S)$	stock adjustment of orders at echelon n
O_n	orders placed by echelon n
D_n	delivery rate at echelon n
A_n	incoming material rate at echelon n
d_n	de-stocking fraction at echelon n
R_n	delivery ratio echelon n
B_n	backlog at echelon n

3.3.2 *Forecasting*

The forecasting sector maintains a sales forecast by accumulating the differences between the incoming customer demand (O_{n-1}) and the previous forecast (F_n). When demand exceeds the forecast it's updated upwards and vice-versa. To allow for a smoothing of the forecast, these differences are divided by the

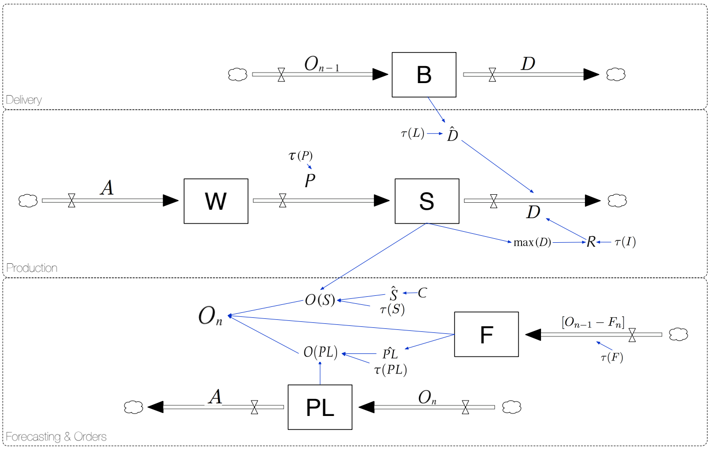


Figure 3.1 *Overview of a modeled echelon.*

forecast adjustment time ($\tau_n(F)$), indicating whether the whole difference or only a fraction is taken into account.

$$(\frac{d}{dt})F_n = \frac{O_{n-1} - F_n}{\tau_n(F)}. \tag{3.1}$$

3.3.3 *Production*

The production sector models the flow of material through the echelon. The incoming material rate (A_n) is equal to the delivery rate of the immediately upstream echelon (D_{n+1}),

$$A_n = D_{n+1}. \tag{3.2}$$

The pipeline inventory is the cumulative difference between orders placed and orders received,

$$(\frac{d}{dt})PL_n = O_n - A_n. \tag{3.3}$$

Incoming material is stored as work in process (W_n). In the interest of simplicity we do not model any production release rule. Thus, the work in process stock is not used strategically or as a control variable: all incoming material is committed to production, and the production rate is modeled by applying a fixed delay (equal to the production time PT_n) to the order arrival rate. System dynamics modeling allows for the introduction of this discrete step in the model, which approximates the real production process,

$$P_n = DELAY(A_n, PT_n). \tag{3.4}$$

Equation 3.4 assumes a production model where the manufacturing time is independent of the utilization rate, it also implicitly assumes that there are no capacity limitations for production (the model can be straightforwardly extended to include capacity limitations).
On -hand inventory (S_n) depends on the delivery rate (D_n) and the production rate (P_n),

$$(\frac{d}{dt})S_n = P_n - D_n. \tag{3.5}$$

Material orders are based on an anchor and adjustment heuristic (Tversky and Kahneman, 1974): the sales forecast acts as the anchor, with the adjustment stemming from the difference between actual and target stock (and supply pipeline) levels.
To calculate the target stock, we start with the desired on hand inventory coverage measured in time units ($\hat{C}_n$). When this is multiplied by the sales forecast, we obtain the desired on hand stock ($\hat{S}_n$) in units of product.

$$\hat{S}_n = \hat{C}_n F_n. \tag{3.6}$$

Analogously, there is a pipeline inventory level $(\hat{PL}_n)$ consisting of the multiplication of the lag (lead time) and the forecasted volumes,

$$\hat{PL}_n = F_n\left(L_n\right). \tag{3.7}$$

3.3.4 *Orders*

Once we have calculated the desired levels of on-hand and pipeline inventories, we generate adjustment orders with the purpose of closing the gap between the actual values of these inventories, and their desired (target) levels. The inventory adjustment time $(\tau_n(S))$ and pipeline inventory adjustment time $(\tau_n(PL))$ represent the time allowed for these quantities to reach the desired levels. These adjustment times model the behavioral aspect of the order generation. Short times imply a nervous buying behavior whereas a long adjusting time is equivalent to a smooth ordering strategy. We define the stock adjustment orders $(O_n(S))$ and pipeline inventory adjustment orders $(O_n(PL))$ as,

$$O_n(S) = \frac{\hat{S}_n - S_n}{\tau_n(S)}, \tag{3.8}$$

$$O_n(PL) = \frac{\hat{PL}_n - PL_n}{\tau_n(PL)}. \tag{3.9}$$

Equations 3.8 and 3.9 calculate the difference between desired and actual values and spread these in equal parts over the amount of periods specified by the adjustment times. Finally, generated orders (O_n) are calculated as,

$$O_n = \max\{0, F_n + O_n(S) + O_n(PL)\}. \tag{3.10}$$

3.3.5 Delivery

A backlog is used to keep track of orders. The backlog is calculated as the cumulative difference between the incoming customer order rate O_{n-1} and actual delivery rate (D_n). O_0, the demand observed by the echelon closest to the end market, is the only exogenous input to the model,

$$(\frac{d}{dt})B_n = O_{n-1} - D_n. \tag{3.11}$$

The order delivery rate (D_n) is the rate of product that is actually shipped out in response to the incoming customer orders. To calculate this, we first define the desired delivery rate ($\hat{D}$), which is equal to the current backlog divided by the expected delivery delay ($\tau_n(L)$),

$$\hat{D}_n = \frac{B_n}{\tau_n(L)}. \tag{3.12}$$

The maximum delivery rate $(max(D)_n)$ per period depends on the ability of firm to physically prepare the products for shipment, modeled as the minimum time to fill orders ($\tau_n(I)$),

$$max(D)_n = \frac{S_n}{\tau_n(I)}. \tag{3.13}$$

We calculate the delivery ratio (R_n) as the proportion of outstanding orders that can be shipped from stock,

$$R_n = \min\{1, \frac{max(D)_n}{\hat{D}_n}\}. \tag{3.14}$$

Finally, the actual order fulfillment rate is equal to the desired delivery rate multiplied by the delivery ratio,

$$D_n = \hat{D}_n R_n \tag{3.15}$$

Alternatively, we can combine equations 3.12 to 3.15 and define the order fulfillment rate as:

$$D_n = \min\{\frac{B_n}{\tau_n(L)}, \frac{S_n}{\tau_n(I)}\}. \tag{3.16}$$

3.3.6 Modeling de-stocking decisions

We model de-stocking decisions by decreasing the desired inventory coverage ($\hat{C}_n$) of an echelon n at time T by a fraction d_n (with $0 \leqslant d_n < 1$).
Thus, we can define $\hat{C}_n$ as:

$$\hat{C}_n = \begin{cases} \hat{C}_n & \text{if } t < T \\ (1-d_n)\hat{C}_n & \text{if } t \geqslant T, \end{cases} \tag{3.17}$$

Where C_n is the desired stock coverage in "normal" (non-crisis) situations. It is important to note that de-stocking is a decision to lower target stock levels that are measured in time units. It is not a decision to reduce its absolute value, nor does it imply the destruction or writing-off of inventory. In this way, we separate explicit decisions to lower inventory targets from the implicit reductions that come from a decrease in sales.

3.3.7 On the equivalence of the ordering policy

The model presented in this section is a straightforward extension of the model found in Sterman (2000). In particular, we introduce an explicit de-stocking decision and a discrete production delay. The structure of this ordering policy, however, is not unique to System Dynamics models, and can be found in other branches of the literature.

In the behavioral operations literature, an equivalent rule is used to model the decision-making behavior of human managers. In this context, the model is presented based upon its equivalence to an 'anchor and adjustment heuristic' (Tversky and Kahneman, 1974). These heuristics are used to describe human decision-making biases: Orders are calculated by selecting an anchor (in this case the forecast) with subsequent adjustments motivated by deviations from the target stock and supply line levels (Sterman, 1989; Croson et al., 2014). In contrast with the models used in laboratory experiments, however, our model allows for dynamic inventory and pipeline coverage (the desired inventory and pipeline are not a constant, but proportional to expected sales) as well as changes in the inventory policies (implemented through the de-stocking hypothesis).

The control theoretic branch of inventory theory also uses a family of models based on the same principles as the model described in this paper. The more general of the models in this framework, the Automatic Pipeline Variable Inventory Order-based Production Control System (APVIOPCS), is a discrete-time, constant-coverage, equivalent of our System Dynamics model. For a discussion on the implications of independent supply line and stock adjustments based upon discrete, control theoretic models, we refer the readers to chapters 4 and 5.

3.4 Results and Analysis

In this section, we use the echelon model as a building block to construct, calibrate, and validate 4 different supply chain models based upon data collected at our research company.

The methodology presented thus far concerns the modeling of a single echelon in a supply chain: The input to an echelon model is a customer order and its output is an order placed to a supplier. To model a supply chain, we link echelon models according to the customer/supplier relationships defined by its structure (e.g. number of echelons, linear, divergent) and parameterize the individual echelons. We run the supply chain models using end-market sales data as their exogenous inputs.

Each of the echelon models is defined by operational and behavioral parameters. Operational parameters posses a concrete interpretation in the day to day operation of a firm (e.g. target stocks and production times) and are thus set based upon expert interviews. Behavioral parameters (e.g. supply line and stock adjustment times), on the other hand, define the relationship between internal variables –product of explicit or implicit managerial decisions– and are thus estimated through a process of model calibration. The de-stocking decisions we hypothesize, however, do not fall squarely in either of these definitions. While these decisions correspond to the operation of the firm, we could find no hard evidence of the desired inventory reductions. Rather, the de-stocking decisions corresponded to financial recommendations from upper management, which were estimated to be 'on the order of 10 to 20 percent'. Similarly, de-stocking decisions do not conform to the definition of a behavioral characteristic of the models. Thus, de-stocking is estimated via scenario analysis based upon expert interviews: Feasible de-stocking quantities (from 5-30% reductions of desired stocks) are defined in discrete increments, the calibration is performed for each of these scenarios, and the best fit is chosen. Potentially, the amount of de-stocking could depend on a series of firm characteristics such as the type of product and distance from the end market. However, due to the limitations of our data, we can only quantify the cumulative effect of de-stocking on the uppermost echelon. We therefore use a single de-stocking parameter for each supply chain.

The rest of this section is divided as follows. We explain the model set-up and data collection in §3.4.1. Then, we define the structure of the modeled supply chains and their observable parameters in §3.4.2, and the calibration of behavioral parameters in §3.4.3. Finally, we study the historical fit of the model in §3.4.4, and analyze an alternative model, where the de-stocking hypothesis is dropped, in § 3.4.5.

3.4.1 *Model set-up and data collection*

Two distinct flows appear when we link individual echelon models to form a supply chain model: an information flow that travels upstream (orders), and a material flow that travels downstream (deliveries). The information flow of any supply chain originates at the sales point of a finished product, i.e., its end-market. Thus, the demand information observed by an upstream firm is a function of the original signal, generated by the end-market, and transformed –throughout its flow upstream– by the subsequent echelons of the particular supply chain (in the case of divergent supply chains, the combination of end-market signals).

We use the 2008 credit crisis as a natural experiment because it allows us to link these end-market signals to the corresponding upstream demand: the synchronization observed during the period effectively controls for the smoothing effects of aggregation. Explicitly, we assume that (a) firms at a given echelon share the same structure, (b) firms at a given echelon share the same behavior during this time frame, and (c) the information distortion that is observed in the passage of demand information upstream –from its origin in the end-market– corresponds to locally rational policies at each stage of the supply chain (i.e., no demand information is arbitrarily created or discarded by intermediate echelons). We base this approach upon the observation that empirical data shows that during the credit crisis turning points in sales and inventories were indeed aligned by tiers (retail, wholesale, and manufacturing) (Dooley et al., 2010), suggesting that companies along a single echelon exhibited a synchronized behavior.

Conceptually, we model echelons that represent individual tiers: a group of competing companies providing the same product to the same supply chain. This follows what Sprague and Wacker (1996) define as the modeling with a "disaggregation by stages along the inventory stream". They point out that management practice generalizations made in this way recognize the impact of the management of inventory as it progresses through the stages. In particular, we model 4 different supply chains that belong to a dutch chemical company. As a reference, the different supply chains belong to four different business units that are situated 4-5 echelons upstream from retail demand. The upstream products of these supply chains are different resins, thermoplastics, and polymers. The end-markets where these are found are shown in Table 3.2

We use monthly, EU27 data available from Eurostat as a proxy for the demand for each of the end-markets for the period of Jan 2005–Aug 2009. The series used are: Construction index, automobile registrations, household goods retail index, and production indexes for: food products (C10), paper and paper products (C17), glass and glass products (C231), basic metal and metal products (C24), and motor vehicles (C291). All data is normalized with the average of $2007 = 100$. The work at each of the four sites of the company began with a kickoff meeting with management where the objectives and scope of the

Table 3.2 *Summary of end-markets served by the 4 modeled supply chains*

Supply Chain	End Markets
Supply Chain 1 (resins a)	Residential and commercial construction; Residential and commercial repair & maintenance.
Supply Chain 2 (resins b)	Residential and commercial construction; Residential and commercial repair & maintenance; Furniture sales.
Supply Chain 3 (polymers)	Automotive sales.
Supply Chain 4 (thermoplastics)	Automotive manufacturing; Glass panel manufacturing; Metal manufacturing; Food manufacturing; Paper and pulp manufacturing; Residential and commercial construction.

study were explained and defined. Following these, interviews were conducted with employees to formalize data collection procedures. The structure of the supply chain model and the parameterization of observable parameters is based on input from these employees, complemented with information obtained from players distributed along the supply chain. The modeling work was performed on-site, which allowed for additional ad-hoc interviews and further familiarization with the particulars of each individual supply chain.

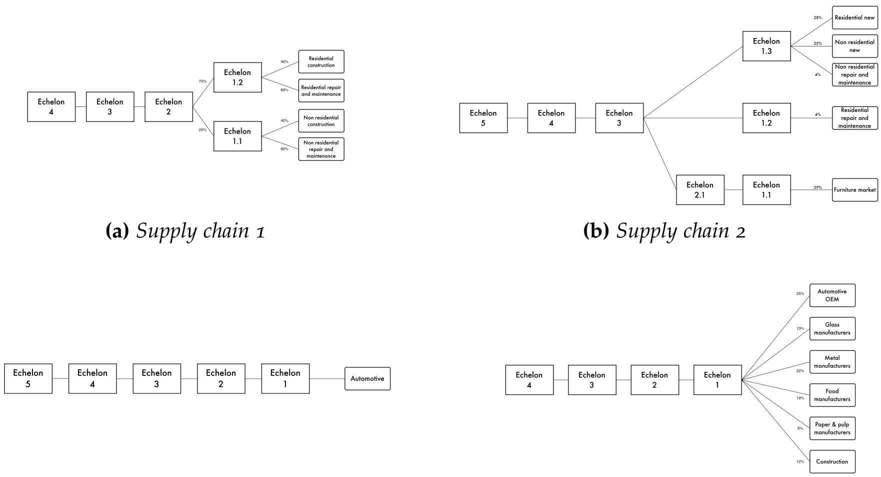
3.4.2 *Structure and observable parameters*

The number of echelons, structure, and end markets of each of the 4 supply chains are all different and can be seen in Figure 3.2. The supply chains in this study consist mainly of chemical firms upstream and make-to-stock component suppliers downstream. For this study, we consider our research site to be the upstream-most boundary of each supply chain. The parameterization of the observable parameters per echelon is shown in Table 3.3.

Table 3.3 *Observable Supply Chain parameters per echelon*

Supply Chain 1	$\hat{C}$	L_n	PT_n	*Supply Chain 2*	$\hat{C}$	L_n	PT_n	*Supply Chain 3*	$\hat{C}$	L_n	PT_n	*Supply Chain 4*	$\hat{C}$	L_n	PT_n
1.1	5	10	1	**1.1**	8	4	1	**1**	4	3	1.5	**1**	8	0.25	2
1.2	5	10	1	**1.2**	4	4	1	**2**	2	1	1	**2**	14	0.25	5
2	8	2	1	**1.3**	4	4	1	**3**	1.5	2	1	**3**	10	0.25	4
3	8	2	1	**2.1**	8	4	1	**4**	1	2	1	**4**	8	0.25	1
4	4	2	1	**3**	8	0.25	1	**5**	2	2	1				
				4	3	0.25	1								
				5	2	0.25	1								
de-stocking	0.15				0.25				0.2				0.1		

In an attempt to simplify the models, we assume deterministic lead times and the availability of resources such that order preparation does not introduce



(a) *Supply chain 1* **(b)** *Supply chain 2*

(c) *Supply chain 3* **(d)** *Supply chain 4*

Figure 3.2 *Supply chain structures*

significant lags. Thus, the expected delivery delay ($\tau_n(L)$) is equal to its own delivery lead time (L_{n-1}), and the minimum time to fill orders ($\tau_n(I)$) is equal to 1. Due to the absence of disaggregated data, the lead time is defined as the time between placing an order and its receipt (i.e., it encompasses both the informational and physical components of the delay). We make two assumptions regarding the boundary conditions: (1) orders placed by the uppermost echelon in a supply chain are always served by a supplier with infinite stock, and (2) downstream demand is exogenous and is composed of the individual demand signals of the end markets that require the materials produced upstream.

With the structure and observable parameters estimated for each of the supply chain models, we proceed with the analysis of the de-stocking decisions and the estimation of behavioral parameters.

3.4.3 *Model calibration and behavioral parameters*

Behavioral parameters are, by definition, not observable; we must estimate them through the individual calibration of each of the supply chain models. Model calibration is the process of estimating parameters to obtain a match between modeled and observed behavior and is, in itself, a stringent test of the validity of the model that links structure and behavior (Oliva, 2003). Nevertheless, Oliva (2003) points out that achieving a good historical fit is not enough to confirm the dynamic hypothesis behind the model; the model has to match the observed behavior for the right reasons.

In system dynamics models, this is usually achieved through "partial model calibration": The process of estimating parameters within a subset of model parameters, instead of the entire model parameter space (Homer, 2012). This strategy "reduces the risk of the structure being forced into fitting the data, increases the efficiency of the estimation (estimators with smaller variances), and concentrates the differences between observed and simulated behavior in the piece of structure responsible for that behavior" (Oliva, 2003). In the context of our supply chain models, partial model calibration entails calibrating each echelon separately. Unfortunately, we cannot do so because we lack primary sales and inventory data at the intermediate levels, and it is not possible to map secondary empirical data to individual echelons.

To overcome this, however, we perform a full model calibration and follow it with: (*i*) a sanity check of the estimated parameters (is the model structure sound?) and (*ii*) the test of an alternative hypothesis (can we achieve the same behavior through a different structure?) to increase the confidence in our model.

We use 27 months (Jan 2007 – Mar 2009) of historical, secondary, EU27 end-market data as the input for each of the end-markets in our models, and primary sales data from our research company as a proxy for upstream demand, which we use as the calibration target.

In using full model calibration, we assume an open-loop behavior for the supply chain models. Namely, we assume that the input to the model (end-market demand) does not depend on its output (upstream demand). This assumption is reasonable because we calibrate an uncapacitated model during a period where demand was either relatively stable or declining, which controls for any upstream influence on downstream demand[1].

The calibration step is implemented within the simulation software (Vensim); simulations are performed for each supply chain by generating model runs where all the observable parameters are fixed as established in §3.4.2, while the behavioral parameters are varied. The cumulative sum of square errors between the estimated demand and the historical sales data is calculated per run and the combination of parameters that minimizes this error is then chosen. Formally,

[1]In a capacitated supply chain, upstream demand can potentially influence downstream demand. Failure to meet upstream demand due to capacity constraints can result in an increase of downstream demand at a later point in time due to, for example, shortage gaming (i.e. when customers, expecting shortages, artificially inflate their orders).

the minimization corresponds to:

$$\min_{\tau_n(F),\tau_n(PL),\tau_n(S)} \sum_{t=1}^{k} (O_{N-1}(t) - \tilde{D}_N(t))^2 \tag{3.18}$$

Where N is the most upstream echelon in the supply chain, and $\tilde{D}_N(t)$ is the historical sales data for time t at echelon N (our research company). In other words, we compare the orders generated by the modeled customers of our research site with the actual historical sales of said firm and search for the parameter values that minimize the error. The minimization is performed through a modified Powell-Brent algorithm (Brent, 2002). For computational purposes and to reduce the search space, $\tau(S)$, $\tau(PL)$, and $\tau(F)$ are estimated through their reciprocals, α_S, α_{PL}, and Θ, and α_S, $\alpha_{PL}, \Theta \in [0,1]$. Table 3.4 lists all the parameters estimated through calibration, including the 95% confidence intervals calculated through a sensitivity analysis. De-stocking fractions represent a non-observable parameter and were thus estimated through a combination of interviews and scenario analysis; experts identified a range of de-stocking fraction per supply chain, scenarios where then run with different values for d (intervals of 0.05), and the best fit was chosen. All echelons in a supply chain are assumed to incur the same de-stocking behavior at the same time. This is motivated by: (a) recent literature suggesting synchronization during the crisis, (b) the availability of only upstream sales data as a calibrating time-series. The latter factor negates the additional information that could be gleaned from having individual de-stocking parameters: upstream sales are affected by the cumulative amount of inventory being taken out of the supply chain. These data are shown in Table 3.3.

In all cases, the confidence bounds of the estimations for the uppermost echelon are lax: This is due to the data available for calibration being the historical sales of this echelon. None of the parameters in the model allow a firm to influence its own demand via strategic decisions. Thus, the uppermost echelon can either meet the demand or incur in destabilizing stock-outs. The confidence bounds represent the parameter space that allows for the former. Similarly, the amount of parameters being estimated from a single time series (between 12 and 21, depending on the supply chain) explain the size of the confidence intervals of the pipeline inventory adjustment time, which are particularly large.

A parameter estimated to be ∞ corresponds to a parameter that is not taken into account in the ordering heuristic. The upper bound for the pipeline inventory adjustment time for all but two of the echelons is ∞, which suggests that we cannot reject the hypothesis that firms completely ignore the pipeline. On the other hand, the lower bound for more than 2/3 of the echelons is larger than one, which suggests smoothing of the pipeline inventory adjustment.

This is consistent with results from experiments found in the behavioral literature (Sterman, 1989). The gap between desired and actual pipelines is severely underestimated in the ordering decisions: Both the means and medians of the pipeline inventory adjustment time ($\tau(PL)$) are larger than the respective values for the stock adjustment times (two-sample t-tests on the means, Wilcoxon rank-sum tests for the medians, all $p \leq 0.01$), as well as the values for the forecast adjustment time ($p \leq 0.1$ when comparing means and $p \leq 0.01$ when comparing medians). When we compare the stock and forecast adjustment times, on the other hand, we find that the mean forecast adjustment time is larger than the mean stock adjustment time ($p \leq 0.1$). However, this difference is caused by extreme values in supply chain 2, as there is no statistical difference between the medians of these parameters—suggesting a smoothing of the same order of magnitude for the adjustment of the inventory gap and the forecast updating. We next compare the adjustment times between the different supply chains to test for any differences in the inherent behavior. We find that the mean and median adjustment times of supply chain 4 are significantly larger than those of supply chain 2 ($p \leq 0.1$). However, all other mean and median differences are non-significant. This suggests that, while supply chain 4 is less responsive than supply chain 2, the overall behavior and the mechanisms behind all the supply chain models is comparable.

3.4.4 *Historical fit and structural validity*

Following the calibration, we run the four supply chain models driven by the exogenous end market and the de-stocking hypothesis. In contrast to the calibration step, runs are now performed with published historical end market data complemented with scenarios based upon published forecasts, and expectations of the business intelligence groups for subsequent periods.
Figure 3.3 shows the model outputs against the seasonally corrected upstream demand realizations. The vertical axis represents the demand expressed in % of the average 2007 demand and the dotted vertical lines indicate the threshold between historical fit and forecast values. Table 3.5 shows the root mean squared error (RMSE), R^2, and Theil inequality statistics for the data series shown in the figure. These inequality statistics decompose the mean square error into three fractions representing: unequal means (U_m), unequal variances (U_s), and imperfect correlation (U_c) (Theil, 1966). A low U_m indicates a strong correspondence between the modeled mean and the actual mean, and a low U_s indicates a similar correspondence between variances. Therefore, low variance and means statistics indicate that the error is unsystematic, and therefore

Table 3.4 *Estimated behavioral Supply Chain parameters*

	$\tau(S)$	95% CI		$\tau(PL)$	95% CI		$\tau(F)$	95% CI	
Supply Chain 1									
1.1	**3.91**	3.14	4.93	**6.16**	4.05	9.48	**5.49**	1.53	11.27
1.2	**8.70**	6.13	15.60	∞	191.79	∞	∞	794. 62	∞
2	**9.70**	8.96	10.70	∞	25.65	∞	**10.30**	8.43	12.76
3	**14.08**	13.17	15.12	∞	31.05	∞	**17.56**	15.54	19.93
4	**10.00**	1.42	∞	**100.00**	1.42	∞	**6.03**	1.00	2194.05
Supply Chain 2									
1.1	**4.15**	3.31	5.25	∞	3.50	∞	**3.91**	1.00	9.85
1.2	**3.99**	1.00	∞	∞	1.00	∞	**4.31**	1.00	∞
1.3	**4.54**	3.09	6.95	∞	5.30	∞	**1282.98**	27.33	∞
2.1	**3.33**	1.18	5.19	**1.00**	1.00	1.76	**4.60**	2.37	7.38
3	**7.20**	6.20	8.4	**10.00**	1.00	∞	**983.40**	157.60	∞
4	**23.11**	17.30	32.31	**3333.33**	1.00	∞	**8.10**	6.30	9.10
5	**10.00**	1.00	102.63	**10.00**	1.00	∞	**6.03**	1.00	33.45
Supply Chain 3									
1	**9.69**	7.22	12.69	∞	7.22	∞	**21.64**	13.69	37.33
2	**6.41**	4.20	8.97	∞	4.20	∞	**16.98**	9.85	30.67
3	**8.74**	6.32	11.81	∞	6.32	∞	**11.02**	6.85	17.08
4	**13.90**	10.02	17.53	∞	10.02	∞	**9.58**	6.61	13.51
5	**10.00**	1.00	∞	**100.00**	1.00	∞	**6.03**	1.00	∞
Supply Chain 4									
1	**10.18**	7.13	14.07	**31.10**	7.13	∞	∞	90.49	∞
2	**9.35**	7.00	12.63	**33.03**	7.00	∞	**16.91**	9.03	31.49
3	**13.76**	10.21	18.94	**2215.27**	10.21	∞	**11.53**	6.83	18.96
4	**16.88**	1.00	∞	**101.73**	1.00	∞	**10.00**	1.00	∞

Table 3.5 *Historical fit statistics*

	RMSE	R^2	U_m	U_s	U_c
Supply chain 1	4.53%	0.65	0.031	0.190	0.779
Supply chain 2	8.11%	0.68	0.065	0.041	0.893
Supply chain 3	6.74%	0.88	0.040	0.210	0.750
Supply chain 4	11.48%	0.75	0.000	0.079	0.921

desirable (Oliva and Sterman, 2001).

The models, driven by one exogenous data series (end customer demand), and one “crisis response” policy (desired stock reductions in September 2008) shows a good tracking of the overall behavior of the system. However, a match between observed and simulated behavior is not in itself enough to accept the model and hypothesis. As Oliva (2003) explains, “There is a chance that a set of parameter values might be capable of replicating the observed behavior through a set of unrealistic formulations, and thus generate the right behavior for the wrong

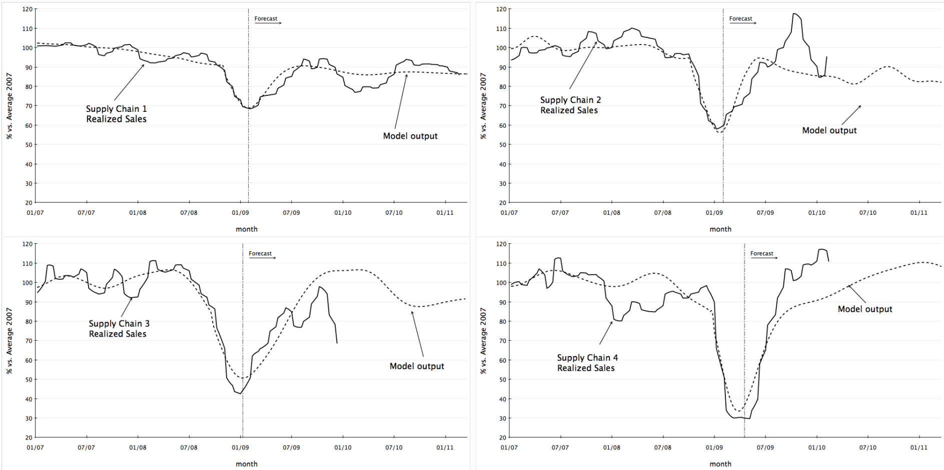


Figure 3.3 *Model output vs. seasonally corrected sales data.*

reasons". To test the validity of the model, we analyze what the estimated parameters say about its structure, and follow with the analysis of an alternative model to test whether the same behavior can be achieved through a different structure.

The system dynamics model presented in §3.3 is based on the anchor and adjustment heuristic for order generation (Tversky and Kahneman, 1974). The same decision rule is used in a substantial amount of experimental work, were individual human behavior is analyzed in the context of the beer distribution game. Three of the most salient such studies are Sterman (1989), Croson et al. (2014), and Croson and Donohue (2006). In these experiments, students (and professionals) play the beer game (under different settings), and then the behavioral parameters are then fit through the use of regression analysis. From these publications, we can derive the estimate of the stock, pipeline inventory, and forecast adjustment times, and we see that in all cases the experimental studies conclude that subjects tend to under-estimate the pipeline inventory, and observe low reactiveness towards inventories and forecasts. The fact that these two phenomena can be observed in the behavioral parameters of our calibrated models increases our confidence on its structural validity.

3.4.5 *Alternative model*

The de-stocking hypothesis presented in this chapter is motivated by a variety of results from the inventory management and economics literature. It has been shown that firms can convert assets into cash in the short-term (Sudarsanam and Lai, 2001), and that lowering inventories is a common response to financial distress (Pesch and Hoberg, 2013). Furthermore, studies focused on the mechanisms behind the recent 2008 financial crisis have reported on its extraordinary magnitude and synchronization (Alessandria et al., 2010). In line with this, anecdotal evidence points to decisions having been made to reduce working capital: formal and informal interviews with decision makers across the industry support this view.

However, the empirical validation of our hypothesis is not possible: inventory targets are not explicitly reported, and we can't use actual inventories as a proxy for inventory targets during the time-frame of this study. The decision to reduce inventory targets triggers a shock that immediately affects orders, but its effect on inventory levels is substantially more complex: the combination of time delays and declining demand caused inventory levels to spike following the start of the crisis, further increasing the gap between target and actual inventory levels.

Therefore, to further test our model and –in particular– the de-stocking hypothesis driving it, we perform additional experiments to rule out alternative explanations. In Figure 3.4 we observe the model output of a calibrated alternative model without the de-stocking hypothesis. The alternative model differs from the original model only in that it does not allow for any de-stocking to be performed at any point in the supply chain, and its calibration follows the same procedure outlined in §3.4.3. Table 3.6 shows the fit statistics.

Table 3.6 *Alternative model fit statistics*

	RMSE	R^2	U_m	U_s	U_c
Supply chain 1	7.69%	0.002	0.004	0.682	0.314
Supply chain 2	11.96%	0.259	0.007	0.260	0.733
Supply chain 3	12.83%	0.611	0.068	0.009	0.923
Supply chain 4	15.96%	0.577	0.011	0	0.988

We see that the alternative model adequately tracks the average, or long-term, demand variations but cannot explain the magnitude of the demand drops nor their timing. If we compare the alternative model runs with the original (de-stocking) model runs we can see that the end market sales drive the long-term evolution of upstream sales, while the short term dynamics are dominated by shocks. In Table 3.7 we present the summary statistics for the relevant variables of both models. We perform two-sample t-tests for the means and

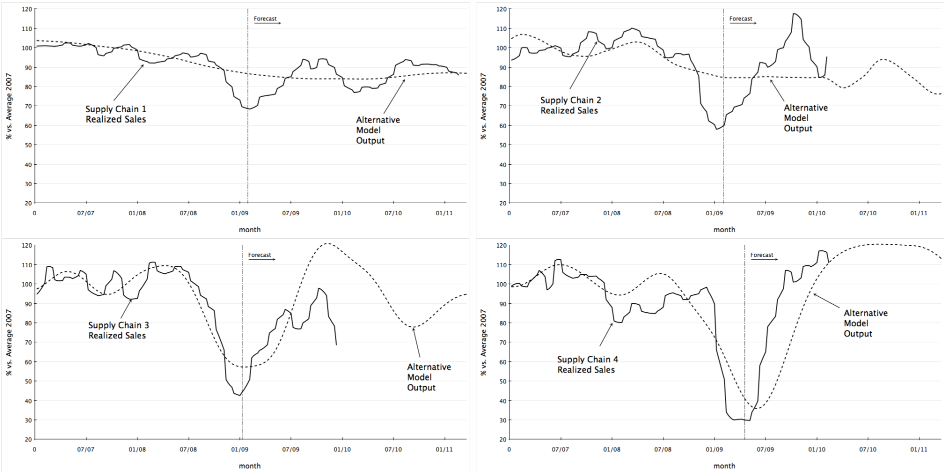


Figure 3.4 *Alternative model output.*

Table 3.7 *Summary statistics for original and alternative models*

Variable	Mean	St. Dev	1st Quartile	Median	3rd Quartile
			Original Model		
$\tau(S)$	9.67	4.81	6.41	9.69	10.18
$\tau(PL)$	7254.17	23152.01	31.10	100	101.73
$\tau(F)$	360.80	1096.79	6.03	10.15	17.27
RMSE	7.72	2.91	5.64	7.43	9.80
R^2	0.74	0.10	0.67	0.72	0.82
			Alternative Model		
$\tau(S)$	10.96	3.51	10	11	12.86
$\tau(PL)$	38009.98	68070.23	17.55	72.72	89756.50
$\tau(F)$	8.89	3.59	6.03	9.96	10.99
RMSE	12.11	3.41	9.83	12.40	14.40
R^2	0.36	0.29	0.13	0.42	0.60

find that the alternative model has a significantly larger RMSE and a lower R^2 than the original model (all $p \leq 0.05$). Similarly, testing the effect on the medians (Wilcoxon rank-sum test) gives the same results ($p \leq 0.1$, and $p \leq 0.05$ respectively).

Additionally, we test for differences in the estimated parameters: the means of the forecasting, and pipeline inventory adjustment times differ at the $p \leq 0.1$ level when comparing original and alternative models, but no such difference is found when testing for difference in the median values. This is consistent with the observation that abnormally large values of these parameters skew the average results. In general, we observe that all the calibrated models exhibit characteristics consistent with experimental findings: slow reaction speeds and an under-weighing of the pipeline inventory.

3.5 Conclusions and Managerial Insights

Behavioral dynamics in supply chains have been widely researched. Initial studies by Forrester (1958) analyzed data at the level of individual or series of companies. Following the work by Lee et al. (1997a), extensive analytical work has been conducted, and more recently, driven by the work by Sterman (1989) and Croson and Donohue (2005), focus has been on laboratory experimentation. On the empirical front, Cachon et al. (2007) does not find conclusive evidence for the existence of the Bullwhip Effect in aggregate empirical data. Chen and Lee (2012), through analytical work, argue that it is the aggregation of the data that plays an important role in hiding some of the effect, which is observed at a firm level (Bray and Mendelson, 2012).

In this study, we use observations following the collapse of Lehman Brothers in the Fall of 2008 to develop a hypothesis regarding target level setting and investigate the explanatory power of behavioral dynamics. Our study observes demand at the level of an individual company, but takes into account the hypothesized dynamic decision making behavior at meso-level. With this, our study sets itself apart from previous studies, and not only builds upon the lines of research discussed above, but also on research in economics studying inventory cycles. Our results show that the theoretical results of Sterman (1989) and Croson and Donohue (2005) together with an endogenous inventory shock, can explain the dynamic evolution of demand observed upstream in the periods following the start of the recent credit crisis. The endogenous replenishment process drives the evolution of demand throughout the supply chain, determined by structural characteristics of the supply chain (following Forrester, 1958) and the hypothesized human behavior (following Sterman, 1989). The empirical evidence presented shows that slow reaction speeds and an apparent underestimation of the supply pipeline are prevalent at higher aggregation levels, suggesting that they go beyond being a phenomenon of individual decision-making biases. At this level, the pipeline underestimation seems to be caused not from an incorrect estimation of target values, but as a combination of the inherent reaction time of firms and a decision rule that eschews the tracking of the pipeline inventory by instead steering on large amounts of on-hand inventory. This finding calls for further study on the ordering behavior of firms; if behavioral biases influence decision-making at the echelon level, *how can –and should– firms overcome them?*. Equally important: *how do these behaviors change over time?*

To increase the confidence in our de-stocking hypothesis in the presence of limited data, we presented an alternative model without the hypothesized de-stocking. Our results in four supply chains show that the underlying behavior is consistent among both models and also with prior research. Furthermore, while the exogenous demand at consumer level and endogenous ordering decisions in the supply chain drive the overall demand evolution, short-term demand bullwhip-like dynamics are mainly driven by the de-stocking response to the crisis.

For managers, our results have implications at both the tactical and strategic levels of decision making. Tactically, for managers it is much more important to keep track of consumer demands, supported by an endogenous simulation of ordering behaviors to make demand forecasts, rather than relying exclusively on information obtained from one or two echelons downstream. These simulation-based forecasts can drive decisions on plant openings and closures, staffing decisions, and aggregate inventory strategies. Additionally, our results highlight the importance of understanding the implications that policy changes can bring into a supply chain. It is well known that aggregate inventory levels can serve as an additional way to achieve liquidity targets, however, limited research exists on the implications of such decisions on the stability of the entire supply chain. Our results suggest that such decisions can be a significant source of demand fluctuations.

Strategically, we show that the structure of the supply chain impacts the clockspeed at which the supply chain operates. In this sense, we provide a formal model that can be used (a) to analyze the effects of structural and policy changes in the supply chain, and (b) to potentially become a decision-making tool in which endogenous behavioral changes form the basis of scenario-based forecasting. A cursory glance at the model outputs of the original and alternative models suggests that in certain cases (such as in supply chain 1), the de-stocking behavior not only had influence in the short term dynamics, but that it is also the main source of the steep demand drops observed in the periods following Lehman Brother's bankruptcy.

In this sense, our findings highlight the prospective value of information sharing. In cases such as the period studied in this paper, and consistent with experimental research (Croson and Donohue, 2006), knowledge about the underlying source of the observed demand dynamics (i.e. distinguishing between 'actual' demand drops and inventory adjustments) is crucial so as to adopt the correct response strategy.

There are limitations and opportunities for further research. First, we used a single time series of upstream sales per echelon model, which hinders our ability to perform partial model calibration and dissociate the de-stocking decisions according to the supply chain stages. To overcome this we performed the model calibration during a time-frame where supply chain decision-making was particularly synchronized and demand was declining. Further studies with detailed data at every stage of the supply chain can offer more robust statistical tests of firm-level behavior as well as bring insights regarding the influence of firm characteristics (in particular the distance from the end market) in the ordering and de-stocking behavior.

Second, we make a series of implicit assumptions that may not necessarily apply in other industries or time periods. Our models assume independent echelons with no information sharing among them, with constant market share (and no price changes), a stable supply chain structure, no capacity limitations, and aggregate data. We expect these assumptions to be reasonable within the crisis time-frame, but further modeling efforts are necessary to test whether they can be applied during stable times, where the demand dynamics are more subtle.

Such studies, combining fine grained data from multiple echelons in a supply chain, have the potential to take us closer to the objective, both empirical and experimental, of testing whether endogenous mechanisms that we know govern the individual behavior (such as the underestimation of the supply line, and de-stocking and hoarding behaviors) can be consistently found at the aggregate level.

I was almost driven to madness in considering and calculating this matter.

Johanes Kepler

Chapter 4
Human Behavior and the Bullwhip: an Analytical model

The Bullwhip Effect is a major problem in today's supply chains. It is a dynamic phenomenon that has sparked a vast body of research from a wide array of methodologies. Empirical, experimental, and analytical studies exist of both a descriptive nature –trying to identify and describe it– and of a normative nature –trying to overcome it–. The causes for the Bullwhip Effect can be broadly separated into operational (such as order batching and price fluctuations) and behavioral categories (such as artificially inflating orders and pipeline underestimation).

In the previous chapter, we saw that a synchronized inventory shock is capable of generating the demand dynamics experienced by upstream manufacturing firms during the 2008 credit crisis. When calibrating that model, we observed that the model's behavioral parameters are consistent with certain behavioral traits that are routinely exhibited by humans in beer game experiments.

The goal of this chapter is to extend the analytical knowledge regarding the influence of human behavior in the appearance of the Bullwhip Effect and the amplification of inventory variance. We use linear control theory as a modeling methodology, and frame our work as a descriptive work: we attempt to link existing experimental and empirical results to insights developed through this analysis.

4.1 *Introduction*

In this chapter, we use discrete-time classical control methods to analyze an Automatic Pipeline, Variable Inventory Order Based Production Control System (APVIOBPCS) design in light that it is the policy that most closely resembles human decision-making heuristics. We aim to understand not only the bias shown by beer game players (Sterman, 1989; Croson and Donohue, 2006)) in under-estimating the pipeline, but also why empirical data suggests that real-life firms operate through APVIOBPCS with low fractional adjustments (see Chapter 3). We focus on both the bullwhip and inventory variance amplification and adopt stationary and transient measures to quantify performance.
The rest of this chapter is structured as follows: In the next section we introduce the discrete-time model in both time and frequency domains. We follow with a comprehensive analysis of the stability of the system, deriving exact expressions for the general stability boundaries. We introduce performance measures in Section 4.4, analyzing the dynamic, and steady state, responses of both the generation of orders and the evolution of inventories. We then identify the trade-offs inherent to these designs, and position real-life firms in this context. We conclude in Section 4.5, and present all the mathematical proofs in Appendix A.

4.2 *Model Description*

In this section, we analyze a discrete-time, periodic-review, single-echelon, general APVIOBPCS design with an exponentially smoothed forecast of demand. The inventory coverage ($C \in \mathbb{R}^+$), the delivery lead time ($L \in \mathbb{N}$), and the forecast smoothing parameter $\alpha \in [0,1]$ are the structural parameters of the system, while the pipeline ($\gamma_P \in [0,1]$) and inventory ($\gamma_I \in [0,1]$) adjustment factors are the behavioral parameters of the system[1]. The inventory coverage represents the target inventory (measured in weeks of sales) that a firm chooses to maintain, while the target pipeline is calculated each period as the product of the forecast and the system's lead time. The lead time is assumed deterministic and defined as the time elapsed between the placement and receipt of a replenishment order. The behavioral parameters specify the fraction of the gap between target and actual values that are taken into account when generating orders: γ_I is the fraction of the inventory gap to be closed per period, and γ_P is the fraction of the pipeline gap to be closed per period. For instance, a system

[1]Note that the APVIOBPCS model presented in this chapter is the discrete equivalent to the model presented in Chapter 3. Because of the different intuitive representation, the behavioral parameters in this chapter are expressed as fractional adjustments, the reciprocal of the adjustment times introduced in Chapter 3.

with $\gamma_I = 1$ and $\gamma_P = 0$ completely closes the inventory gap every period, while it ignores the pipeline entirely.

Formally, the sequence of events, and the equations in the model are as follows: at the beginning of each period (t) a replenishment order (o_t) based on the previous period's demand forecast (f_{t-1}) is placed with the supplier. Following this, the orders that were placed L periods before are received. Next, the demand for the period (d_t) is observed and served. Excess demand is back-ordered. Then, the demand forecast is updated according to the formula: $f_t = \alpha d_t + (1-\alpha) f_{t-1}$. The forecast is used to compute the target levels of both inventory, $\hat{i}_t = C f_t$, and pipeline, $\hat{p}_t = L f_t$. The orders that will be placed in the following period (o_{t+1}) are generated according to an anchor and adjustment-type procedure, $o_{t+1} = \gamma_I(\hat{i}_t - i_t) + \gamma_P(\hat{p}_t - p_t) + f_t$. The balance equations for inventory (i) and pipeline (p) are: $i_t = i_{t-1} + o_{t-l} - d_t$, and $p_t = p_{t-1} + o_t - o_{t-l}$. Note that the assumptions that orders and inventories can be negative are necessary to maintain the linearity of the model.

This sequence of events is identical to the one described in Hoberg et al. (2007b) with the difference being that our study introduces the fractional behavioral parameters γ_I and γ_P. Other studies of AP(V)IOBPCS designs use different order of events; e.g., Dejonckheere et al. (2003), and Disney (2008) orders are placed at the end of each period. These changes in the sequence of events introduce extra unit delays in the equations. However, these differences only affect the mathematical representation of the system; the structure of the system and the results, remain the same.

4.2.1 *On the equivalency of APVIOBPCS and Order Up To policies*

Dejonckheere et al. (2003) remark that when $\gamma_I = \gamma_P = 1$, an APVIOBPCS design is equivalent to an Order-Up-To (OUT) policy with safety lead times. To see this equivalency, we take a classical OUT policy, where orders are defined as follows:

$$o_t = \hat{D}_t^L + SS_t - IP_t, \tag{4.1}$$

where $\hat{D}_t^L$ is the expected lead time demand, SS_t the safety stock, and IP_t the inventory position. The safety stock is given by $SS_t = k\hat{\sigma}_t^L$, where k is a constant term chosen to meet a required service level, and $\hat{\sigma}_t^L$ is the expected standard deviation of the forecast error during the lead time. Following Disney et al. (2006b) we can set the safety constant $k = 0$ and instead set a safety lead time, C, so that we can rewrite the policy (using the notation and sequence of events defined in the prior section) as:

$$o_{t+1} = \hat{D}_t^{L+1} + \hat{D}_t^C - IP_t. \tag{4.2}$$

Since the inventory position is equal to the sum of net and pipeline stocks, and the expected demand is maintained in our model through the forecast, f_t, we can successively obtain:

$$o_{t+1} = \hat{D}_t^{L+1} + \hat{D}_t^C - I_t - p_t, \tag{4.3}$$

$$o_{t+1} = (L+1)f_t + Cf_t - I_t - p_t, \tag{4.4}$$

$$o_{t+1} = f_t + Cf_t - I_t + Lf_t - p_t. \tag{4.5}$$

The latter equation represents an APVIOBPCS design with $\gamma_I = \gamma_P = 1$.

Following this reasoning, an APVIOBPCS design with independent adjustment of the inventory deficits through $\gamma_I \neq \gamma_P$, is a generalization of OUT policies; analogous to a Proportional-Order-Up-To (POUT) policy (see Chen and Disney, 2007) with individual adjustments of the net and pipeline inventory deficits:

$$o_{t+1} = f_t + \gamma_I(Cf_t - I_t) + \gamma_P(Lf_t - p_t). \tag{4.6}$$

If we restrict $\gamma_I = \gamma_P = \gamma$ we obtain an analogous of a POUT policy where the adjustments are driven by the inventory position deficit:

$$o_{t+1} = f_t + \gamma\left((L+C)f_t - IP_t\right). \tag{4.7}$$

Given that the safety stock in APVIOBPCS designs is determined by the inventory coverage C, the trade off between cost and service level considerations is implicitly included in its calculation, as well as in the smoothing parameters γ_I and γ_P.

When we bound the fractional adjustments between zero and unity ($\gamma_I, \gamma_P \in [0, 1)$), these are smoothing policies. While classical OUT policies react swiftly to changes in the demand at the expense of variability in orders, smoothing policies display steadier orders at the expense of a slower reaction time—they use inventories as a buffer for demand variability. All else equal, a smoothing policy will result in higher inventory-related costs to achieve the same service level as a classical OUT policy[2]. From a cost perspective, then, the adoption of smoothing APVIOBPCS policies hinges on the trade-off between the costs related to the variability of inventory and orders, and the cost of inventory and holding costs.

Because the work contained in this chapter is descriptive in nature, we assume

[2] This observation is trivial if we consider that it has been shown that OUT policies minimize inventory and shortage costs (Dejonckheere et al., 2003).

that the parameters C, γ_I, and γ_P are given and therefore we do not attempt to calculate or optimize them. In the remainder of this chapter, we explore the influence of different parameters in the dynamic performance of APVIOBPCS policies without explicit considerations of the cost trade-offs associated.

4.2.2 *The Frequency Domain*

The model we introduced completely describes the relationships between the parameters of a general APVIOBPCS design. However, due to time dependencies, we cannot find a clear relationship between the inputs and the outputs of the system. For this reason, we turn from the time domain to the frequency domain (where these relationships become simply algebraic), by taking the Z-transform of the system's set of equations.

The Z-transform is defined as:

$$\mathbb{Z}\{x_t\} = X(z) = \sum_{k=0}^{\infty} x_k z^{-k}, \tag{4.8}$$

where z is a complex variable and x_k is the value of a time series in period k. We refer the reader to Jury (1964), and Nise (2007) for a comprehensive background on discrete systems and the Z-transform method; and to Hoberg et al. (2007b), and Dejonckheere et al. (2003) for an introduction to their application on inventory modeling.

Using the following properties of the Z-transform:

$$\mathbb{Z}\{a_1 x_t + a_2 y_t\} = a_1 X(z) + a_2 Y(z) \text{ (Linearity)}, \tag{4.9}$$

$$\mathbb{Z}\{x_{t-T}\} = z^{-T} X(z) \qquad \text{(Time delay)}, \tag{4.10}$$

we can write all system parameters in the frequency domain. The equation for orders is written then as:

$$O(z) = \frac{\left[\gamma_I(\hat{I}(z) - I(z)) + \gamma_P(\hat{P}(z) - P(z)) + F(z)\right]}{z}. \tag{4.11}$$

In control theory, the response of a system is completely characterized by its transfer function $G(z) = N(z)/C(z)$, that represents the change in output $N(z)$ with regards to a change in input $C(z)$ in the frequency domain. In this chapter we are interested in studying the properties of the transfer function of orders ($G_O(z)$) and the transfer function of inventories ($G_I(z)$) The transfer function of orders represents the change in orders $O(z)$ in response to a change in customer demand $D(z)$:

$$G_O(z) = \frac{O(z)}{D(z)} = \frac{\left[\gamma_I(\hat{I}(z) - I(z)) + \gamma_P(\hat{P}(z) - P(z)) + F(z)\right]\frac{1}{z}}{D(z)}$$
$$= \frac{\left[\alpha(\gamma_I C + \gamma_P L + 1)(z-1) + \gamma_I(z-1+\alpha)\right] z^L}{(z-1+\alpha)(z^L(z-1+\gamma_P) + (\gamma_I - \gamma_P))}. \quad (4.12)$$

Analogously, the transfer function of the inventory is defined as the change in inventory level $I(z)$ as a response to customer demand $D(z)$:

$$G_I(z) = \frac{I(z)}{D(z)} = \frac{\frac{z}{z-1}\left[O(z)z^{-L} - D(z)\right]}{D(z)}$$
$$= \frac{z\alpha(\gamma_I C + \gamma_P L + 1)(z-1) + z(z-1+\alpha)\left(\gamma_P - z^L(z-1+\gamma_P)\right)}{(z-1)(z-1+\alpha)(z^L(z-1+\gamma_P) + (\gamma_I - \gamma_P))}. \quad (4.13)$$

Having defined the transfer functions for orders and inventories[3], we can now provide structural properties of the response of the system.

4.3 *Stability and Aperiodicity of the System*

We have the following definition for the stability of a system:

Definition 4.1 *(Nise, 2007). A system is stable if every bounded input yields a bounded output, and unstable if at least one bounded input yields an unbounded output.*

In our model, the customer demand is the input, and orders and inventory are the outputs. In this context, a stable production/inventory system will produce finite orders and finite inventories as long as the demand itself is finite. Thus, the stability of the system is a desirable property that we want to study. However, Definition 4.1 does not specify mathematical conditions for stability, necessary to understand whether a given system is stable or not. An alternative definition-condition that connects the stability of a discrete system with its transfer function is:

Definition 4.2 *(Jury, 1964). Suppose that $G(z) = N(z)/C(z)$ is the transfer function of a linear time-invariant, system and that the denominator $C(z)$ has exactly n roots p_i, namely $C(p_i) = 0$, $i = 1, \ldots, n$. We call the roots p_i poles of the transfer function, and we say that a system is stable if all poles p_i are within the unit circle of the complex*

[3]Readers familiar with the stability criteria for a discrete system will notice that the existence of the term $(z-1)$ in the denominator of Eq (4.13) suggests that the system is not stable. We cope with this issue in §4.3.1, following the formal definition of the stability criteria.

plane ($|p_i| < 1$)*, marginally stable if at least one pole is on the unit circle* ($|p_i| = 1$)*, and unstable if at least one pole resides outside the unit circle* ($|p_i| > 1$)*.*

Consequently, judging the stability of a system is equivalent to finding the solutions to the characteristic equation $C(z) = 0$.

Remark 4.1 *Suppose that P with* $|P| \geq 1$ *is a root of* $C(z)$ *with multiplicity m, namely,* $C^{(i)}(P) = 0, \forall\, i = 0, \ldots, m-1$*. If* $N^{(i)}(P) = 0, \forall\, i = 0, \ldots, m-1$*, and if all the other roots of C(z) are inside the unit circle, then the system is called stabilizable. However this is sometimes used alternatively as a definition for a stable system (Wunsch, 1983). This is not the case here.*

In the next section we derive conditions for the stability of the system through an analysis of the structure of the involved characteristic polynomials, and we introduce the aperiodicity of the system, which is a characterization of the dynamic response of a stable system. We begin our analysis with the response of orders to changes in demand (Eq. (4.12)) and follow with the analysis of the inventory response to changes in demand (Eq. (4.13)).
In the forthcoming mathematical analysis, the fractional parameters γ_I, and γ_P are not a-priori bounded to $[0, 1]$. This allows us to characterize the stability of the system over a broader range of possible values than the ones that can be found in a physical system.

4.3.1 *Stability Boundaries*

By comparing equations (4.12) and (4.13) we see that the characteristic polynomials of orders and inventories are almost equal except for the extra term $(z-1)$ that appears in the latter. The pole $z = 1$ would render the inventory response marginally unstable, unless this is also a root of the numerator of $G_I(z)$ (see Eq. (4.13) and Remark 4.1). To this effect, we use the geometric series $z^{L+1} - 1 = (z-1)\sum_{i=0}^{L} z^i$ to rewrite equation (4.13) as:

$$G_I(z) = \frac{z\alpha(\gamma_I C + \gamma_S L + 1) - z^{L+2} - z^{L+1}(\alpha + \gamma_P - 1) + \gamma_P z + \alpha\gamma_P \left(1 - \sum_{i=0}^{L} z^i\right)}{(z-1+\alpha)(z^L(z-1+\gamma_P) + (\gamma_I - \gamma_P))}. \tag{4.14}$$

Thus, in an APVIOBPCS design, the stability of both orders and inventories is defined by the same characteristic polynomial

$$C(z) = (z-1+\alpha)(z^L(z-1+\gamma_P) + (\gamma_I - \gamma_P)), \tag{4.15}$$

Being a polynomial in z of degree $L+2$ with real coefficients, $C(z)$ has exactly $L+2$ roots. Unfortunately, this polynomial is transcendental: it is impossible to find its roots independently of L. Furthermore, exact solutions for $C(z)=0$ can only be found for values of $L \leq 2$. Thus, we study structural properties of $C(z)$ to derive a set of conditions that define an exact stability boundary for the general APVIOBPCS design[4]. The proofs of all the theorems, lemmas, and propositions are found in the Appendix.
It can be shown that APVIOBPCS designs share the same characteristic polynomial with APIOBPCS designs (Disney and Towill, 2006). The latter have constant target inventories as opposed to being proportional to expected sales as in APVIOBPCS. Thus, the insights and conclusions derived from the analysis of C(z) hold for both designs. For notational simplicity we refer to AP(V)IOBPCS designs when both can be used interchangeably.

Proposition 4.1 *The stability of a general AP(V)IOBPCS system with smoothing parameter $\alpha \in [0,1]$ can be determined by analyzing the poles of the reduced characteristic polynomial:*

$$\hat{C}(z) = z^L(z-1+\gamma_P) + (\gamma_I - \gamma_P). \tag{4.16}$$

An AP(V)IOBPCS is stable if all the roots of $\hat{C}(z)$ are located inside the unit circle.

Thus, the stability of a general AP(V)IOBPCS system with the commonly used exponential smoothing parameter range of $[0,1]$ is completely determined by the values of L, γ_I, and γ_P.

Theorem 4.1 *For each value of L, stability is guaranteed when γ_I and γ_P satisfy the following $L+1$ conditions:.*

(i)

$$|\gamma_I - \gamma_P| < 1, \tag{4.17}$$

(ii)

$$(1-(\gamma_I-\gamma_P)^2)\,|\gamma_P - 1|^{(n-1)}\,U_{n-1}(X) - |\gamma_P - 1|^n\,U_{n-2}(X) > 0, \quad n = 2,\cdots,L, \tag{4.18}$$

[4]There are simplified criteria to determine the stability of a system without having to explicitly calculate the roots of the characteristic polynomial: The Routh-Hurwitz stability criterion (Nise, 2007) for continuous systems and its discrete analogous, Jury's stability criterion (Jury, 1964). However, these methods require the knowledge of the order of the characteristic polynomial. Because the order of the characteristic polynomial in our system is defined by the lead time L, such methods cannot be used to obtain a stability criterion independent of L.

where $U_n(X)$ is the Chevyshev polynomial of the second kind, defined by

$$U_n(X) = \frac{\left(X + \sqrt{X^2 - 1}\right)^{n+1} - \left(X - \sqrt{X^2 - 1}\right)^{n+1}}{2\sqrt{X^2 - 1}}, \tag{4.19}$$

with

$$X = \frac{1 - (\gamma_I - \gamma_P)^2 + (\gamma_P - 1)^2}{2\,|\gamma_P - 1|}, \tag{4.20}$$

and,

(iii)

$$\left(\left(1 - (\gamma_I - \gamma_P)^2\right)^2 - \left((\gamma_I - \gamma_P)\,(\gamma_P - 1)\right)^2\right) |\gamma_P - 1|^{L-1}\, U_{L-1}(X) -$$
$$- 2\left(1 - (\gamma_I - \gamma_P)^2\right) |\gamma_P - 1|^{L}\, U_{L-2}(X) + |\gamma_P - 1|^{L+1}\, U_{L-3}(X)$$
$$+ 2\,(-1)^{L+1}\,(\gamma_I - \gamma_P)\,(\gamma_P - 1)^{L+1} > 0. \tag{4.21}$$

Remark 4.2 *It can be seen that the L conditions defined by* (4.17) *and* (4.18) *describe regions of convergence decreasing in L. These regions are plotted in Figure 4.1. We observe that the intersection of all the regions that are defined by the conditions in* (4.18) *is equal to the region that is defined in* (4.18) *for $n = L$. Moreover, we found that the last condition can be simplified. This allows us to pose the following conjecture.*

Conjecture 4.1 *For each value of L, stability is guaranteed when γ_I and γ_P satisfy the following conditions:*

(i) $|\gamma_I - \gamma_P| < 1$,

(ii) $(1 - (\gamma_I - \gamma_P)^2)\,|\gamma_P - 1|^{(L-1)}\, U_{L-1}(X) - |\gamma_P - 1|^{L}\, U_{L-2}(X) > 0$,

(iii) If the L is odd, then the third condition reduces to $\gamma_I > \max\{0, 2(\gamma_P - 1)\}$*, and if L is even, then the third condition reduces to* $0 < \gamma_I < 2$.

This conjecture has been numerically verified for $L = 2, \cdots, 200$. Furthermore, the behavior of the conditions point towards an asymptotic region of convergence, defined by

Lemma 4.1 *For all values of lead times $L \in \mathbb{N}$, stability is guaranteed in the area that is bounded by the lines $\gamma_I = 0$, $\gamma_I = 2$, $\gamma_I = 2(\gamma_P - 1)$, and $\gamma_I = 2\gamma_P$.*

To build intuition for the reasoning behind Conjecture 4.1 and Lemma 4.1, we refer the reader to Figure 4.1 where we plot the (γ_I, γ_P) pairs that satisfy the conditions from Theorem 4.1 in the form of colored regions in the $[\gamma_I, \gamma_P]$ space for six different values of L. We plot the region that satisfies condition *(i)* in a shade of blue, the region that satisfies condition *(iii)* in a shade of light red, and the $L-1$ regions that satisfy condition *(ii)* in shades of yellow. A system satisfies all conditions, and thus is stable, if its pair of parameters (γ_I, γ_P) falls inside a region where all $L+1$ conditions are satisfied. We color the regions that satisfy each of the conditions with semitransparent hues so that graphically, the intersection of all the $L+1$ regions (which defines the stability of the system) is represented by a dark purple stability region (the variation in the color shades across plots corresponds to the amount of overlapping regions, which increases in L). The additional green areas in Figs. 4.1a and 4.1b concern aperiodicity, which we define in the next sub-section.

The plots on the left column of Fig. 4.1 correspond to systems with odd lead times, and the plots on the right column correspond to even lead times. It is easy to see that the region that satisfies condition *(i)* is independent of the lead time, whereas the region that satisfies condition *(iii)* depends on the parity of the lead time. Indeed, we observe that the latter region is the same for all even (odd) lead times, this observation results in the simplified condition *(iii)* defined in Conjecture 4.1.

When we analyze the behavior of the $(L-1)$ regions that satisfy condition *(ii)*, we observe that within the overlap of the regions that already satisfy conditions *(i)* and *(iii)* (red and blue), each additional yellow region includes the preceding one. In other words, the n^{th} yellow region lies inside the $(n-1^{st})$ yellow region. This observation results in the simplified condition *(ii)* defined in Conjecture 4.1. Thus, we can ensure stability by determining the overlap of only 3 regions for any value of lead time L.
Furthermore, looking at the region of stability, we can see that it asymptotically converges towards the parallelogram defined in Lemma 4.1.

Note that in this mathematical analysis we have not bounded the parameters γ_I nor γ_P. From our analysis, this results in pairs of values where for a given γ_I, there may exist a $\gamma_P < 0$ for which the system is stable. Having $\gamma_P < 0$ implies that the adjustment of orders will increase when the pipeline gap decreases and vice versa. Disney (2008) described this "negative" stability region for a system with $L = 2$ and found that its response is heavily dampened. Similarly, for all examples shown in Figure 4.1, there exists a region of stability where $\gamma_I > 1$ and/or $\gamma_P > 1$. In this case, the order adjustments calculated by such a policy will be larger than the inventory/pipeline gap. Given that such policies are counter intuitive, and not analogous to commonly used OUT policies, we restrict the analysis in the remainder of this chapter to the case where $\gamma_I, \gamma_P \in [0, 1]$.

4.3.2 Aperiodicity

If a system has a time-domain response with a number of maxima or minima that is less than n, the order of the system, we call such a system aperiodic (Jury, 1985). These dynamics are also defined by the poles of the transfer function: positive real poles contribute a damping component to the response, whereas negative real poles, and poles with an imaginary component will contribute oscillatory terms (Nise, 2007). Formally,

Definition 4.3 *(Jury, 1985). Suppose that $G(z) = N(z)/C(z)$ is the transfer function of a stable, linear, time-invariant, system. Thus, all poles of the transfer function, p_i, $i = 1, \ldots, n$ are within the unit circle. The response of this system is aperiodic if $\forall i$, $p_i \in [0, 1)$. From Disney (2008) we adopt the concept of a weakly aperiodic system if $\forall\, i, p_i \in \mathbb{R}$ and there exists an index $k \in \{1, \ldots, n\}$ such that $p_k < 0$.*

By analyzing the poles of the reduced characteristic polynomial (4.16) for AP(V)IOBPCS and applying Definition 4.3 we obtain the following propositions:

Proposition 4.2 *When $\gamma_I = \gamma_P = \gamma$ the response of a stable system for all lead times L is:*

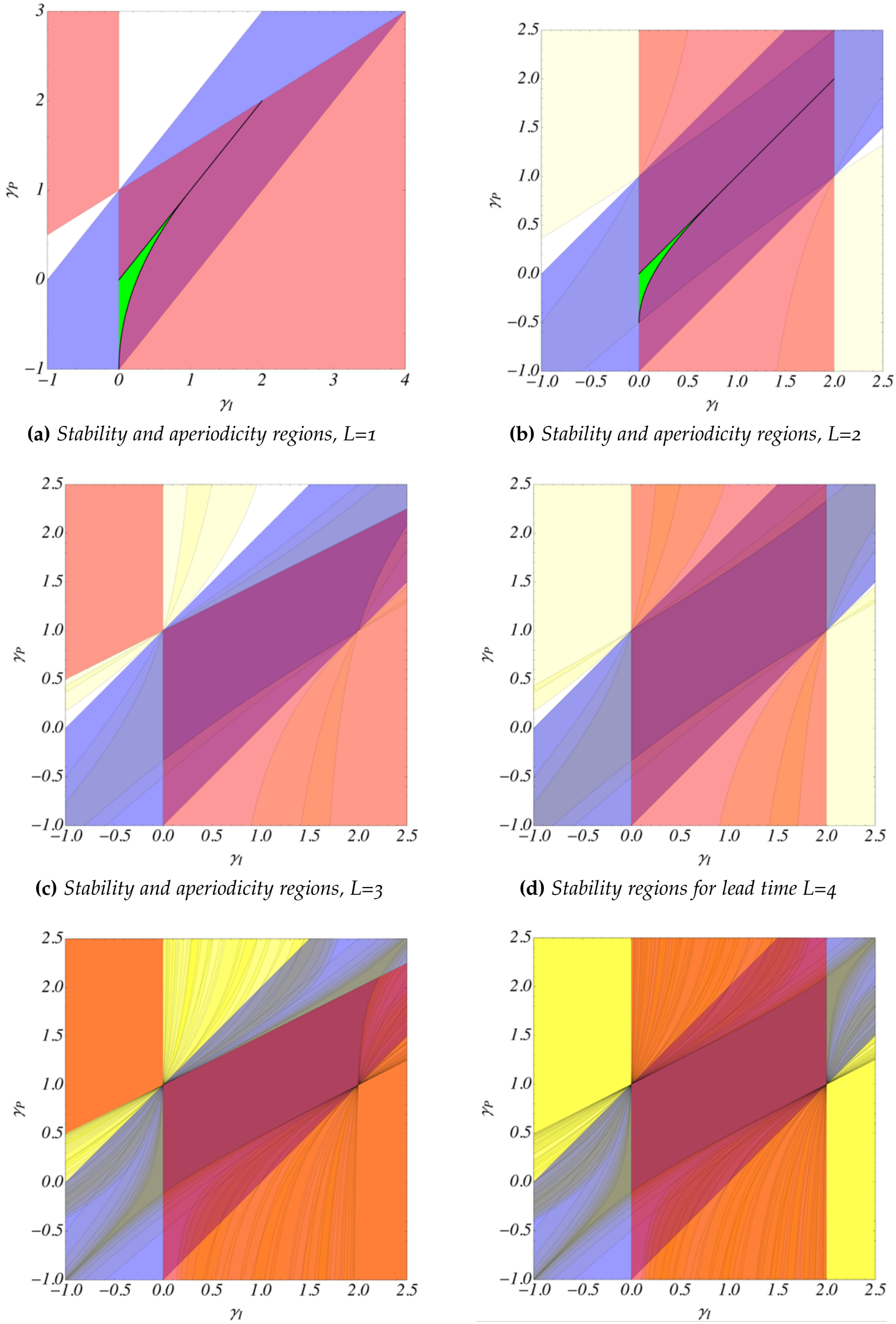
- *aperiodic when $0 < \gamma \leq 1$, and*
- *weakly-aperiodic when $1 < \gamma < 2$.*

Proposition 4.3 *When $L > 2$ and $\gamma_I \neq \gamma_P$ the response of a stable system is non-aperiodic.*

We can find aperiodicity and weak-aperiodicity for the cases of $\gamma_I \neq \gamma_P$ and $L = 1, 2$. The area shaded in green represents the region for which the system is aperiodic in Figure 4.1a and weakly-aperiodic in Figure 4.1b. The boundaries for these regions can be found by following the same analysis as in the proof of Proposition 4.3.

The analysis of stability of an AP(V)IOBPCS design is important because a stable system guarantees bounded orders and inventories for any possible demand, as long as it is finite. Similarly, the aperiodicity analysis of the system is relevant because an aperiodic system avoids costly oscillations. By themselves, however, stability and aperiodicity boundaries are not enough to measure

the performance of the system. The stability conditions and aperiodicity propositions, as well as the special regions defined in the accompanying figures, must be seen, then, as a necessary first step in the evaluation of the system. The analysis presented thus far is necessary, but it is not enough to distinguish any performance difference between two stable systems, or two aperiodic systems.
From an analysis of the poles of the transfer functions, we see that policies where $\gamma_I < \gamma_P$ will always have a dampening component, whereas when $\gamma_I > \gamma_P$ the response can be oscillatory. This observation suggests that the performance of the system will depend on the ratio of the behavioral parameters; we expect an oscillatory response in systems where the pipeline is under-estimated ($\gamma_I > \gamma_P$), and an over-dampened response in systems where the inventory is under-dampened. However, we cannot derive more general statements on the performance of the system through pole analysis because the amount, and magnitude of the poles (necessary to characterize the response) depend on the order of the system, and on the behavioral parameters. This, coupled with the observation made, where humans tend to under-estimate the pipeline, motivate the structure of the next section. We aim to characterize the performance of a stable system as a function of the ratio of its behavioral parameters through extensive numerical experimentation.



(a) *Stability and aperiodicity regions, L=1*

(b) *Stability and aperiodicity regions, L=2*

(c) *Stability and aperiodicity regions, L=3*

(d) *Stability regions for lead time L=4*

(e) *Stability regions for lead time L=9*

(f) *Stability regions for lead time L=10*

Figure 4.1 *Stability and aperiodicity conditions.*

4.4 The relationship Between Behavior and Performance

In this section, we study the performance of an APVIOBPCS design through an analysis of its variance amplification (stationary performance) and its response to demand shocks (transient performance). We introduce in each case the relevant measures we need to characterize its performance. Our objective is to gain an understanding of the influence of the behavioral parameters on said performance measures through extensive experimentation.

4.4.1 Variance Amplification

In supply chains, the concept of variance amplification is often studied in the context of the Bullwhip Effect: the propensity of orders to be more variable than demand signals, and for this variability to increase the further upstream a firm is in a supply chain (Lee et al., 1997a). In general, variance amplification measures the ratio of input variance to output variance and can be defined for any pair of input/outputs. In this section, we use concepts from control theory to analyze the stationary performance of an APVIOBPCS design through the lens of variance amplification for orders and inventories.
To perform our analysis, we introduce two performance measures: the amplification ratio, which measures the ratio of the output and input standard deviations in the steady state; and the bullwhip measure, which measures the ratio of the output variance to the input variance when demand is stochastic and stationary.

Steady state performance

We study the systems' steady state performance by evaluating its amplification ratio when demand is a sinusoid of frequency ω. In steady state, a sinusoidal input to a linear system produces a sinusoidal output of the same frequency but of a different magnitude and phase. For a given linear system, the ratio between the amplitude of the input and the magnitude of the output at a given frequency is constant and is calculated as the modulus of its transfer function evaluated at that frequency (Dejonckheere et al., 2003). Thus, the steady state amplification ratio of an APVIOBPCS design can be calculated directly, for any input sinusoid, from its transfer functions. Furthermore, it can be shown that for sinusoidal inputs, *the amplification ratio value is exactly the same as the ratio of the standard deviations of input over output* (Jakšič and Rusjan, 2008). Formally, for our system we define $A_{O,\omega}$ as the amplification ratio of orders for a sinusoidal demand of frequency ω, and $A_{I,\omega}$ as the amplification ratio of inventory for a sinusoidal demand of frequency ω, where:

$$A_{O,\omega} = |G_O(e^{i\omega})|, \text{ and} \tag{4.22}$$

$$A_{I,\omega} = |G_I(e^{i\omega})|. \tag{4.23}$$

Frequency response plots are the graphical representation of the $A_{O,\omega}$ and $A_{I,\omega}$ of the system for sinusoidal demands with frequencies between 0 and π. Because any demand stream can be decomposed into a sum of sinusoids, these plots are a very powerful tool; by showing the response to every possible demand frequency, they provide a good intuition over the performance of a certain behavioral policy (pairs of γ_I and γ_P values) with regards to any possible demand pattern.

Stationary performance

We now introduce the bullwhip measure as a way of evaluating the performance of a system over all possible demand frequencies. Disney and Towill (2003) define the bullwhip measure as the ratio between input variance to output variance ($BW = \sigma^2_{out}/\sigma^2_{in}$) and show that if the mean of the input is zero and its variance is unity, the bullwhip of a system can be directly calculated through the square of the area below the frequency response plots. In particular, this implies that if the input to our system is a stationary i.i.d. normal demand stream, then the bullwhip of orders can be calculated through

$$BW_O = \frac{\sigma^2_O}{\sigma^2_D} = \frac{1}{\pi}\int_0^{\pi} |G_O(e^{i\omega})|^2 d\omega, \tag{4.24}$$

and the bullwhip of inventories (BW_I) is defined analogously:

$$BW_I = \frac{\sigma^2_I}{\sigma^2_D} = \frac{1}{\pi}\int_0^{\pi} |G_I(e^{i\omega})|^2 d\omega. \tag{4.25}$$

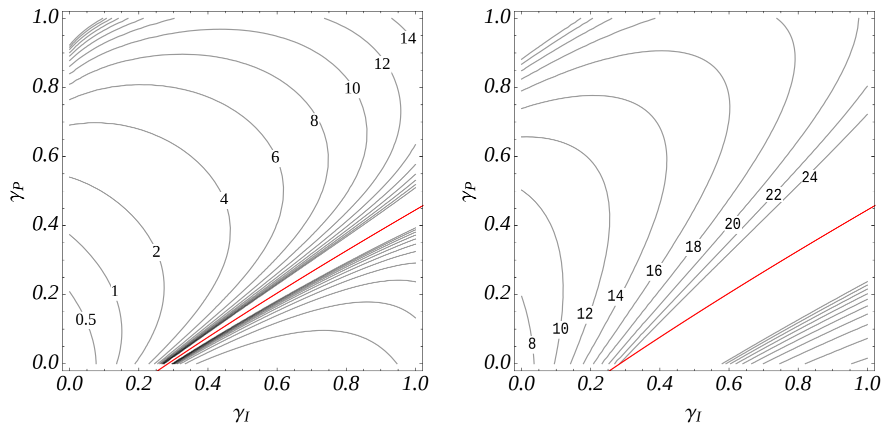
Thus, by plotting the contours of BW_O and BW_I as a function of the behavioral parameters γ_I and γ_P, we obtain a graphical representation of the stationary performance of the system as a function of the behavioral policies.

4.4.2 *Influence of Behavioral Policies on the Amplification of Variance*

Since the stationary and steady state performance measures are closely related, we adopt a two-step approach to analyze the influence of different behavioral policies on the system's performance: we first look at the stationary performance by plotting contours of the bullwhip measures, and follow by analyzing the steady state performance of different policies belonging to the same contour.

Because every frequency is equally represented in the stationary measures, the insights obtained by comparing behavioral policies within a single contour, hold for every contour. The results presented in this section correspond to a system with $\alpha = 0.3, C = 3, L = 5$. Note that we performed extensive experiments with different parameters, we chose to present only these cases since the qualitative conclusions are similar.

Bullwhip contours Figure 4.2a shows the bullwhip contour plots for orders as a function of the behavioral parameters, and Figure 4.2b, the equivalent plot for inventories. We see that despite differences in the magnitude of the variance amplification (BW_I>BW_O for any given behavioral policy), low values of γ_I and γ_P correlate with low values of BW_I and BW_O. Apart from this tendency to increase from the lower left quadrant towards the upper right, the variance amplification displays an asymptote in the critical stability line. These

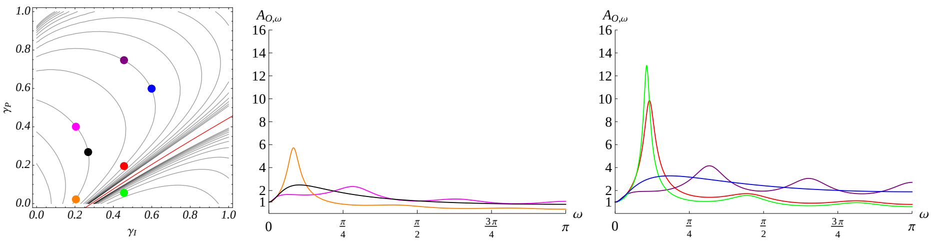


(a) *Order variance amplification (BW_O)* **(b)** *Inventory variance amplification (BW_I)*

Figure 4.2 *Contour plots*

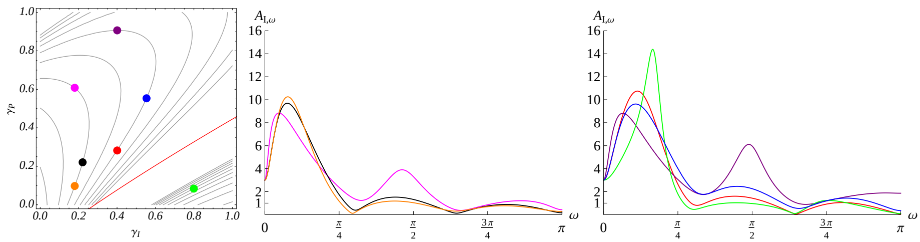
observations suggest that the behavioral policies observed in experimental and empirical research result in good stationary performance for both inventories and orders, but do not, however, provide any additional information about how specific policies affect the performance (i.e., we cannot separate between different policies within a single contour line). Also note that under stationary demand an unstable system has a finite bullwhip (as evidenced by the existence of contour lines in the lower right corner of the graphs).

Performance within contours We use frequency response plots of different behavioral policies within a single contour to study how these policies affect the performance of a system.



(a) *Parameter settings* **(b)** *Order amplification* $BW_O = 2$ **(c)** *Order amplification* $BW_O = 6$

Figure 4.3 *Bullwhip contour, and frequency plots for orders*



(a) *Parameter settings* **(b)** *Inventory amplification* $BW_I = 12$ **(c)** *Inventory amplification* $BW_I = 16$

Figure 4.4 *Inventory amplification contour, and inventory frequency plots*

The steady state performance of the system can be grouped into three categories, according to whether the ratio between behavioral parameters is greater, smaller, or equal to one. Specifically, $\gamma_I > \gamma_P$ policies outperform the rest when the frequency is larger than roughly $\pi/4$ (If demand is observed daily, a frequency of $\pi/4$ represents a cyclical demand with a period of 8 days), but significantly under-performs otherwise. The performance advantages observed are a lower amplification ratio for any given frequency than all other policies, and a higher robustness to changes in frequency (i.e., flatter response) than $\gamma_I < \gamma_P$ policies. Conversely, $\gamma_I < \gamma_P$ policies offer a performance advantage with frequencies smaller than roughly $\pi/4$. We see this in Figure 4.3 (Figure 4.4), where we plot the frequency responses of the same system under 7 different behavioral policies, grouped into Figures 4.3b and 4.3c (4.4b and 4.4c) according to their BW_O (BW_I). We see that the influence of behavioral parameters on the system's performance depends on the demand pattern. In particular, the behavioral policies cause systems to react very differently towards high- and low-frequency demands, with a very clear trade-off in performance. The observed empirical behavior of under-estimating the pipeline, for example, is consistent with a desire to buffer short term fluctuations in demand. In the next section, we study a special case of demand: a one-time shock.

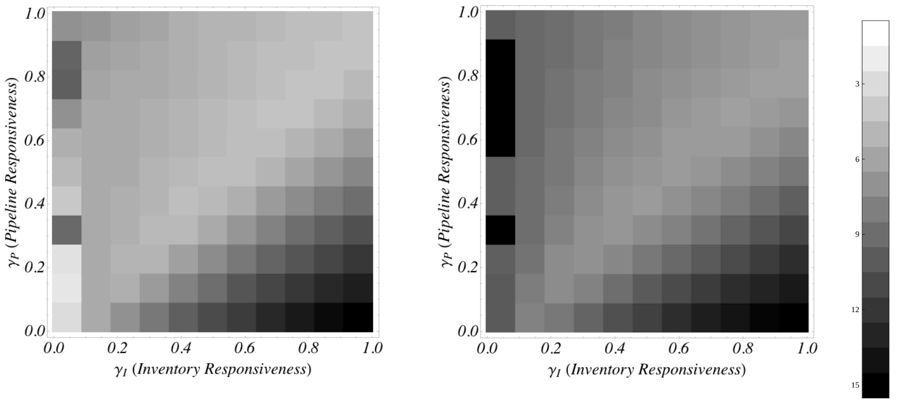
4.4.3 *Influence of Behavioral Policies on the Response to Demand Shocks*

We perform this study in the time-domain, by keeping track of the actual changes in orders and inventories after a step demand increase. We quantify the system's transient performance through the integral time weighted absolute error (*ITAE*), a measure of the dynamic performance of the system in terms of time-weighted deviations from the ideal response (Hoberg et al., 2007b). We use the ITAE to quantify the transient behavior of the system after a step change in demand. The ITAE is defined as:

$$\text{ITAE} = \sum_{t=0}^{\infty} t|\epsilon(t)|, \tag{4.26}$$

where $\epsilon(t)$ represents the absolute error at time t. This measure penalizes deviations from the new steady state, and introduces a linear penalty for longer lasting deviations. Thus, both the amplification and the settling time of the system play a role in its determination. Due to the transcendental nature of the transfer functions of the system, it is not possible to derive a general, closed form expression for its ITAE. An analytical expression for fixed values of L can be calculated through a double transform of the transfer functions (Hoberg et al., 2007b), but the complexity of the general transforms makes this procedure unsuitable for anything but trivial values of L. Hence, we continue to build upon our results thus far through numerical experimentation.

Thus far, the influence of behavioral parameters on the orders and inventories have been comparable. When responding to shocks, on the other hand, we find that $\gamma_P = \gamma_I$ policies perform best when looking at $ITAE_O$, while the performance of $ITAE_I$ is best for a policy of the type $\gamma_P < \gamma_I$. Furthermore, and in direct contrast with the findings of the previous section, for a given γ_I/γ_P ratio, the system's performance increases towards the top-right quadrant of the behavioral parameter space. This means that the behavioral policies that were found to work best for stationary demands, are at a disadvantage when shocks occur. Figure 4.5 shows the transient performance of a APVIOBPCS design as measured by the ITAE of orders ($ITAE_O$) and inventory ($ITAE_I$) of a design with



(a) *Log of ITAE for orders* **(b)** *Log of ITAE for inventories* **(c)** *Log scale*

Figure 4.5 *ITAE as a response to a step increase in demand.*

$\alpha = 0.3$, $C = 3$, and $L = 5$ for different behavioral parameters (γ_I and γ_P). The ITAE is calculated for the first 50 periods. Due to the extreme variation in values, we plot the logarithm of ITAE and clip the values of exploding series.

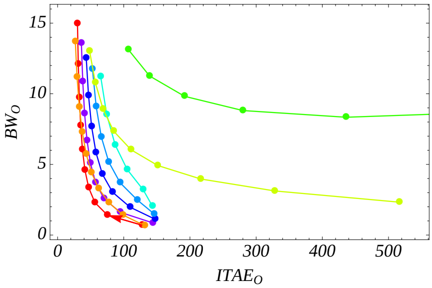
4.4.4 *Insights*

Systems in real life seem to operate in what we can describe as the lower left quadrant of the behavioral parameter space, and below the stability diagonal (Sterman (1989), Chapter 3 of this dissertation). These systems offer performance advantages under stationary demands, and demands with high frequency components but at the same time offer performance disadvantages when there is a demand shock. This is to be expected: the slow behavioral response that buffers high frequency demand fluctuations causes the system to be too slow to respond to a shock, while the rapid behavioral response that allows the system to quickly adapt to a demand shock causes the system to overreact when demand is stationary or cyclic. To understand and analyze this performance trade-off we introduce performance trade-off curves for orders and inventories.

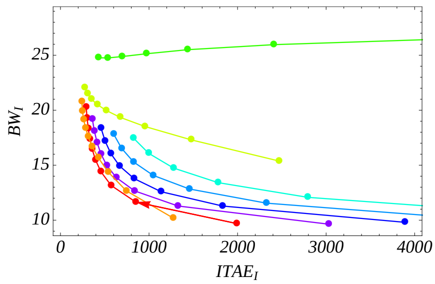
4.4.5 *Stationary and Transient Performance Trade-off*

To quantify the trade-off we group behavioral policies through the γ_I/γ_P ratio (diagonals in the behavioral parameter space) and plot BW_O against $IATE_O$, and BW_I against $IATE_I$. As expected, DE-policies ($\gamma_P = \gamma_I$) offer the best trade-off between stationary and transient performance. However, under-estimating the pipeline ($\gamma_P < \gamma_I$) can give a performance edge if the objective is to minimize the stationary error. This slight performance edge comes at the cost of an increased sensitivity: a deviation from the optimal policy brings about significant performance penalties. Over-estimating the pipeline, on the other hand, decreases the performance of the system, but displays increased robustness.

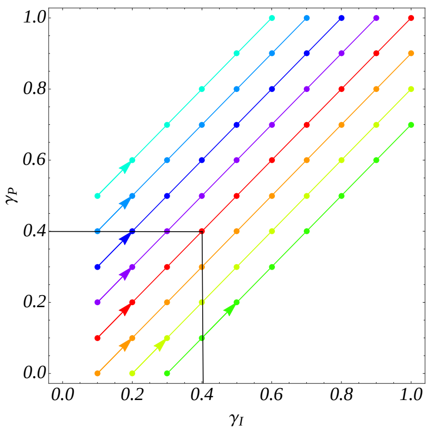
Figures 4.6a and 4.6b show the performance trade-off curves of order and inventories, Figure 4.6c gives the reference for the position of the curves within the behavioral parameters, and Figures 4.6d and 4.6e show the trade-off curves limited to the $[0, 0.4]^2$ area in the behavioral space.



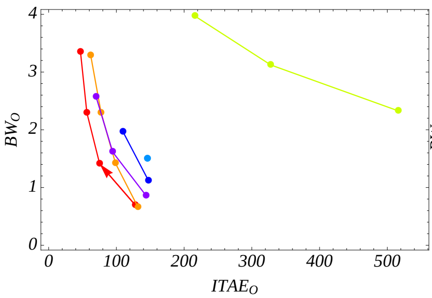
(a) *Order performance trade-off curves*



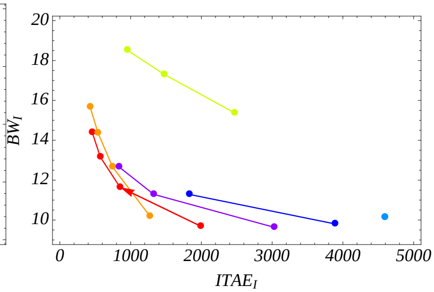
(b) *Inventory performance trade-off curves*



(c) *Reference of diagonals in the parameter space*



(d) *Order performance trade-off curves*



(e) *Inventory performance trade-off curves*

Figure 4.6 *Performance trade-off curves between stationary and transient responses.*

4.5 Conclusions

We have used classical control theory to model general AP(V)IOBPCS systems and analyze the impact of fractional behavioral parameters γ_I and γ_P on their performance. Behavioral parameters in the context of these systems are what in the control theory world are known as feedback controllers, and represent the incomplete closure of inventory and pipeline gaps at the moment of generating replenishment orders. The study of the behavioral influence is motivated by the existence of a large body of experimental, and empirical, research that identifies human biases in ordering decisions. The quantification of these biases suggests that both individual decision makers and firms operate in a clearly defined zone within the parameter space. We attempt, throughout this chapter, to understand what –if any– advantages this zone brings to the performance of the system.

To achieve this, we first modeled a discrete time, general APVIOBPCS design with independent behavioral parameters. We then derived its order and inventory transfer functions, analyzed the stability of such designs and developed a new procedure for the determination of the exact stability region of a system with a given lead time. This test has the advantage of avoiding the direct calculation of determinants, or matrix-based procedures that characterize previous exact solutions of the problem (Jury, 1964; Disney, 2008). Additionally, this procedure allowed us to find an asymptotic region of stability, providing us with a sufficient condition for stability that is independent of L. From the work of Disney (2008), we adopted a characterization of the response of the system based on the position of the poles of the transfer function in the complex plane with which we completely identified the regions for which a system complies with the aperiodicity and weak aperiodicity conditions.

Following the stability and aperiodicity results we performed an extensive set of numerical experiments to help us understand the influence of behavioral parameters in the performance of the system. Due to the large amount of parameters of the system, we needed to specify values for non-behavioral parameters in our experiments. We then performed extensive tests that suggest that the insights developed through the chapter hold in general.

The performance of a system was measured first in the frequency domain, and then in the time domain. Through the frequency domain analysis we found insights related to the steady state response of the system to cyclical inputs (i.e., sinusoidal demands of varying frequency), and the stationary response to white noise (i.e., i.i.d normally distributed demands). Through these analyses, we found that the heavy smoothing (low behavioral parameters) and pipeline underestimation (i.e., not taking the pipeline into account as much as actual

inventory when making replenishment decisions; $\gamma_P < \gamma_I$) found in practice favors the reduction of both the bullwhip for orders and the amplification of inventory variance, as well as increasing the robustness of the system to short term demand fluctuations.
In contrast, the time domain analysis identifies serious performance issues of the system when confronted to a demand shock. The smoothing that we found beneficial in lowering the variance amplification of the system, lowers the reaction time of the system thus decreasing its reactiveness in reaching a new steady state after a shock. Similarly, the pipeline underestimation that contributes to the robustness of the system to short term demand variations, causes performance-decreasing oscillations following a demand shock.

Finally, we showed the trade-off between the stationary and transient performances and found that both inventories and orders achieve good performance around the same values of behavioral parameters. The best combined performance occurs in the lower left quadrant of the behavioral parameter space, which suggests that the smoothing observed in real life data is indeed beneficial. With regards to the pipeline underestimation, also predominant in the data, further research needs to be undertaken to further quantify the trade-offs, taking into account realistic demand streams and cost considerations. From a managerial perspective, these factors have important consequences; the optimal policy will ultimately depend on the cost structure (whether inventory-related, or variability-related costs dominate) as well as on the demand characteristics (whether demand is relatively stationary or not) of a given firm.

Since demand observed in real life is neither purely stationary nor composed entirely of shocks, how should firms approach the trade-offs? What is the weight that should be placed upon different demand types? Are the benefits of increased high frequency robustness achieved through underestimating the pipeline offset by the decrease in performance following shocks? We are confident that our modeling framework provides a good basis to address these and similar highly relevant research questions.

Appendix A
Proofs

Proof of proposition 4.1:
We denote each root of the characteristic polynomial (pole of the transfer function) by p_i^L, where $i = \{1, 2, ..., L+2\}$.
It follows from (4.15) that $p_1^L = (1-\alpha)$ is a root of the characteristic polynomial that does not depend on L.
When $\alpha \in (0,1]$, p_1^L is inside the unit circle and the remaining $L+1$ roots of the characteristic polynomial $C(z)$ will be equal to the $L+1$ roots of the reduced characteristic polynomial $\hat{C}(z)$. Thus the condition for stability when $\alpha \in (0,1]$ reduces to checking that all roots of $\hat{C}(z)$ be inside the unit circle.
When $\alpha = 0$, then $\left|p_1^L\right| = 1$, which means that the system will be marginally stable unless $\left|p_1^L\right| = 1$ is also a root of the numerator of the transfer function. The transfer function for orders when $\alpha = 0$ can be rewritten as:

$$G_O(z) = \frac{\gamma_I (z-1) z^L}{(z-1)C\hat{(z)}} = \frac{\gamma_I z^L}{C\hat{(z)}}. \tag{A.1}$$

The transfer function for inventories when $\alpha = 0$ can be rewritten as:

$$G_I(z) = \frac{\gamma_I - z(z^{L+1} + z^L(\gamma_P - 1) - \gamma_P}{(z-1)C\hat{(z)}} = \frac{\gamma_I - z(z^L + \gamma_P z^{L-1} + \ldots + \gamma_P z + \gamma_P)}{C\hat{(z)}}. \tag{A.2}$$

Thus, since $z = 1$ is a root of the denominator of both $G_O(z)$ and $G_I(z)$, $\forall L \in \mathbb{N}$, the conditions for stability when $\alpha = 0$ reduce to checking that all roots of $\hat{C}(z)$ be inside the unit circle. □

Proof of proposition 4.2:

When $\gamma_I = \gamma_P = \gamma$ we can rewrite the reduced characteristic polynomial:

$$\hat{C}(z) = z^L(z - 1 + \gamma). \tag{A.3}$$

Its $L + 1$ zeroes are:

$$p_2^L = (1 - \gamma), \tag{A.4}$$

$$p_3^L = p_4^L = ... = p_{L+2}^L = 0. \tag{A.5}$$

When $\gamma \in (0, 2)$, $|p_2^L| < 1$, therefore the system is stable. More precisely, for $\gamma \in (0, 1]$, $p_2^L \geq 0$, and for $\gamma \in (1, 2)$, $p_2^L < 0$. Thus, the system is respectively aperiodic, and weakly aperiodic. □

Proof of proposition 4.3:
We know that all the roots of the reduced characteristic polynomial $\hat{C}(z)$ of a stable system lie inside the unit circle of the complex plane. To judge on the aperiodicity of such a system, we need to know whether the roots of $\hat{C}(z)$ lie on the negative or positive half plane.
For this reason, we apply Descartes' rule of signs to the polynomial $\hat{C}(z)$. Descartes' rule of signs states that:

> When the terms of a single variable polynomial with real coefficients are ordered by descending variable exponent, then the number of positive roots of the polynomial is either equal to the number of sign differences between consecutive nonzero coefficients, or is less than it by a multiple of 2. As a corollary, the number of negative roots is the number of sign changes after multiplying the coefficients of odd powers by (-1), or less than it by a multiple of 2. Finally, for a polynomial of degree n, the minimum number of complex roots equals $n - (p + q)$ where p is the maximum number of positive roots, and q the maximum number of negative roots. (Struik, 1969)

To apply Descartes' rule of signs, we write the reduced characteristic polynomial as

$$\hat{C}(z) = z^{L+1} - z^L(1 - \gamma_P) + \gamma_I - \gamma_P. \tag{A.6}$$

We assume that the system is stable and distinguish between two cases: $\gamma_I < \gamma_P$, and $\gamma_I > \gamma_P$.

The case of $\gamma_I < \gamma_P$ For all values of L, γ_P this polynomial will have one sign change. Therefore we will always have one positive and real root. To find the negative and real roots of $\hat{C}(z)$, we separate between odd and even lead times L: For L odd, the polynomial $\hat{C}(-z) = z^{L+1} + z^L(1-\gamma_P) + \gamma_I - \gamma_P$ has 1 sign change for all values of γ_P . Therefore, there exists a real and negative root of $\hat{C}(z)$ and the remaining $L-1$ roots come in pairs of complex conjugates. Only for $L = 1$ does $\hat{C}(z)$ have no complex roots. For L even, the polynomial $\hat{C}(-z) = -z^{L+1} - z^L(1-\gamma_P) + \gamma_I - \gamma_P$ does not have any sign change when $\gamma_P \in [0,1]$ and thus no negative and real roots. This means that it has at least 1 pair of conjugate complex roots. When $\gamma_P \in (1,2)$, it has 2 sign changes and thus 0 or 2 negative and real roots. If it has 2 negative roots and $L = 2$, then the system is weakly aperiodic. In any other case, there exists at least 1 pair of complex roots and the system is thus unstable.

The case of $\gamma_I > \gamma_P$ For all values of L, and for $\gamma_P \in [1,2)$ the reduced characteristic polynomial $\hat{C}(z)$ will have no sign changes and consequently it does not have any positive and real roots. To find the negative and real roots of $\hat{C}(z)$ with $\gamma_P \in [1,2)$, we separate between odd and even lead times L: For L odd, the polynomial $\hat{C}(-z)$ has 2 sign changes and therefore either 2 or 0 real and negative roots. When $L = 1$ and it has 2 negative real roots, the system is weakly aperiodic. In all other cases, there exist at least 1 pair of complex roots and the system will consequently be non-aperiodic. For L even, the polynomial $\hat{C}(-z)$ has 1 sign change and therefore 1 real and negative root. Thus in this case the polynomial will always have at least 1 pair of complex roots and the system will consequently be non-aperiodic. For all values of L, and for $\gamma_P \in [0,1)$ the reduced characteristic polynomial $\hat{C}(z)$ will have 2 sign changes and consequently it has either 2 or 0 positive and real roots. To find the negative and real roots of $\hat{C}(z)$ with $\gamma_P \in [0,1)$, we separate once more between odd and even lead times L: For L odd, the polynomial $\hat{C}(-z)$ has 0 sign changes and therefore either 0 real and negative roots. When $L = 1$ and it has 2 positive real roots, the system is aperiodic. In all other cases there exist at least 1 pair of complex roots and the system will consequently be non-aperiodic. For L even, the polynomial $\hat{C}(-z)$ has 1 sign change and therefore 1 real and negative root. Only the combination $L = 2$ and 2 positive real roots gives a weakly aperiodic response. In all other cases, there exists at least 1 pair of complex roots and thus the system is non-aperiodic. □

Proof of Theorem 4.1:
According to Theorem 43.1 of Marden (1969), the number of roots of our reduced characteristic polynomial $\hat{C}(z)$ (Eq. 4.16) inside the unit circle is equal to the

number of negative signs in the sequence

$$\Delta_1, \frac{\Delta_2}{\Delta_1}, \ldots, \frac{\Delta_{L+1}}{\Delta_L}, \tag{A.7}$$

where

$$\Delta_n := \det \begin{bmatrix} A_n & A_n^{*T} \\ A_n^* & A_n^T \end{bmatrix}, \tag{A.8}$$

and

$$A_n := \begin{bmatrix} a & 0 & 0 & \cdots & 0 \\ 0 & a & 0 & \cdots & 0 \\ 0 & 0 & a & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a \end{bmatrix}, n = 1, \ldots, L, \quad A_{L+1} := \begin{bmatrix} a & 0 & 0 & \cdots & 0 \\ 0 & a & 0 & \cdots & 0 \\ 0 & 0 & a & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b & 0 & 0 & \cdots & a \end{bmatrix},$$

$$A_n^* := \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ b & 1 & 0 & \cdots & 0 \\ 0 & b & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}, n = 1, \ldots, L+1,$$

with $a = \gamma_I - \gamma_P$, and $b = \gamma_P - 1$. Here A_n^T denotes the transpose of A_n where the dimension of these matrices is $n \times n$.

To guarantee stability, we need all the roots of Eq. (4.16) to be inside the unit circle. Thus, we need to have $L+1$ negative signs in the sequence (A.7). Consequently, we need to have:

$$(-1)^n \Delta_n > 0, \quad \forall n = 1, \ldots, L+1. \tag{A.9}$$

Since the matrices A_n and A_n^* commute, according to Silvester (2000), we have that for
$n = 1, \ldots, L+1$, $(-1)^n \Delta_n \; = \det \left(A_n A_n^T - A_n^* A_n^{*T} \right)$. Thus, for $n = 1, \ldots, L$,

$$(-1)^n \Delta_n = \det \begin{bmatrix} 1-a^2 & b & 0 & \cdots & \cdots & -ab \\ b & 1-a^2+b^2 & b & 0 & \cdots & 0 \\ 0 & b & 1-a^2+b^2 & b & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b & 1-a^2+b^2 & b \\ -ab & 0 & \cdots & 0 & b & 1-a^2+b^2 \end{bmatrix}, \tag{A.10}$$

and also ,

$$(-1)^{L+1} \Delta_{L+1} = \det \begin{bmatrix} 1-a^2 & b & 0 & \cdots & \cdots & -ab \\ b & 1-a^2+b^2 & b & 0 & \cdots & 0 \\ 0 & b & 1-a^2+b^2 & b & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b & 1-a^2+b^2 & b \\ -ab & 0 & \cdots & 0 & b & 1-a^2 \end{bmatrix}. \tag{A.11}$$

If we denote M_n as the square $n \times n$ matrix with diagonal elements equal to $1 - a^2 + b^2$, and all the elements on the upper and lower diagonal are equal to b, then we can find recursively that,

$$(-1)^n \Delta_n = (1 - a^2) \det M_{n-1} - b^2 \det M_{n-2}, \quad n = 2, \ldots, L. \tag{A.12}$$

The determinant $\det M_n$ can be calculated through formula (3) of Marr and Vineyard (1988) as

$$\det M_n = D_n(1 - a^2 + b^2, b, b) = |b| \, U_n \left(\frac{1 - a^2 + b^2}{2\,|b|} \right), \tag{A.13}$$

where U_n is the nth degree Chebyshev polynomial of the second kind, defined by

$$U_n(Z) = \frac{\left(Z + \sqrt{Z^2 - 1}\right)^{n+1} - \left(Z - \sqrt{Z^2 - 1}\right)^{n+1}}{2\sqrt{Z^2 - 1}}. \tag{A.14}$$

If we set $X = (1 - a^2 + b^2)/(2\,|b|)$, then

$$(-1)^n \Delta_n = (1 - a^2)\, |b|^{(n-1)}\, U_{n-1}(X) - |b|^n \, U_{n-2}(X). \tag{A.15}$$

Similarly, for the $(L+1)st$ determinant, it holds that

$$
\begin{aligned}
(-1)^{L+1}\Delta_{L+1} =& (1-a^2)^2 \det M_{L-1} - 2(1-a^2)b^2 \det M_{L-2} + b^4 \det M_{L-3} + 2(-1)^{L+1}ab^{L+1} - (ab)^2 \det M_{L-1} \\
=& \left(\left(1-(\gamma_I-\gamma_P)^2\right)^2 - \left((\gamma_I-\gamma_P)(\gamma_P-1)\right)^2\right)|\gamma_P-1|^{L-1} U_{L-1}(X) - \\
& - 2\left(1-(\gamma_I-\gamma_P)^2\right)|\gamma_P-1|^{L} U_{L-2}(X) + |\gamma_P-1|^{L+1} U_{L-3}(X) \\
& + 2(-1)^{L+1}(\gamma_I-\gamma_P)(\gamma_P-1)^{L+1}. \qquad \text{(A.16)}
\end{aligned}
$$

Finally, observe that $-\Delta_1 = 1-a^2$, which completes the proof. $\square$

Proof of Lemma 4.1:
In a compact form, the region defined by Lemma 4.1 can be written as $|b| \leq 1-|a|$ with $a = \gamma_I - \gamma_P$, and $b = \gamma_P - 1$. We observe that condition (iii) of Conjecture 4.1 always defines two boundary lines in this region. Therefore, inside this region, condition (iii) is always going to be satisfied. In order to show that this is indeed the asymptotic region defined by Lemma 4.1 when L goes to infinity, it is sufficient to show that when we set $|b| = 1-|a|$,

$$\lim_{L\to+\infty}(-1)^L\Delta_L = 0. \qquad \text{(A.17)}$$

Knowing that $X = 1$ (see proof of Theorem 4.1) and, using that $U_L(1) = L+1$ (Abramovich and Stegun, 1965, Table 22.3.7), we can rewrite (A.15) as:

$$(-1)^L\Delta_L = (1-|a|)^L(1+L\,|a|), \qquad \text{(A.18)}$$

which goes to 0 as L goes to ∞ since $|a| < 1$ and the proof is complete. $\square$

An experiment is a question which science poses to Nature, and a measurement is the recording of Nature's answer.

Max Planck

Chapter 5
Seasonality and the Bullwhip: Numerical Experiments

In the previous chapter we introduced the frequency-domain analysis of a production/inventory system through the use of linear control theory. We obtained exact expressions for the stability of the system and analyzed the trade-offs inherent to the smoothing of orders. In this chapter, we use linear control theory in a different way. By means of extensive numerical experimentation, we capitalize on the benefits of frequency-domain analysis in the study of cyclical inputs and outputs. In operations management, cyclical inputs (outputs) are of the utmost importance and often found in demand (order) time series—we call them *seasonal* demands (orders).

In this chapter, we use the frequency domain to characterize the seasonality of the orders generated by an APVIOBPCS design. We find that the behavioral parameters not only affect the performance of the system as measured by the metrics introduced in Chapter 4, but that they can also potentially introduce a seasonal component, in both orders and inventories, that is completely unrelated to the seasonality of the demand stream. We observe that inventories are particularly sensitive to these effects.

We find that this phenomenon (identified as rogue seasonality, or fake business cycle) has been a focal point of Forrester's (1958) seminal work, but has since then been seldom analyzed in a systematic way. We review the existing rogue seasonality literature and propose an extension to a methodology to quantify the phenomenon. We show, through numerical experiments, that rogue seasonality information is not contained in neither the stationary nor the transient performance metrics presented thus far. Rather, the rogue seasonality index quantifies information contained in the frequency response plots.

5.1 Introduction: The Bullwhip Effect & Rogue Seasonality

With the publication of his seminal work, Forrester (1958) pioneered a way of modeling and analyzing complex, feedback-ridden systems that would kick-start the field of system dynamics. He analyses the behavior of entire production/inventory systems facing different demand streams and identifies the way in which, even in the absence of changes of both average and periodic retail levels, the system converts random retail-level fluctuations into upswings and downswings in orders and inventories at all levels.

This tendency for production/inventory systems to generate and amplify demand fluctuations has been intensely studied in the years since, and plays a fundamental role in explaining the Bullwhip Effect. The study of the bullwhip has produced a vast amount of analytical, experimental, and empirical research—a summary of which is presented in Chapter 2.

In this chapter, we focus on one of the fundamental concepts related to the Bullwhip Effect that has least been studied over the years: rogue seasonality. This concept, also known as fake business cycle, is clearly illustrated in Forrester (1958), where the author shows –for the production/inventory system– how the upswings and downswings generated by random demand fluctuations can become cyclical—leading to oscillations with a frequency reflecting the characteristics of the system itself rather than external phenomena. In other words, the system generates cyclical orders, and inventories, even in the presence of a random demand.

In recent years, this dynamic component of the Bullwhip Effect has been the subject of a number of studies over which we intend to build upon. Kim and Springer (2008) and Springer and Kim (2010) use a system dynamics model to derive analytic conditions for the appearance of rogue seasonality. They find that the existence of this effect depends on the lead time and the fractional adjustments of inventory and pipeline. However, to keep the models tractable, they approximate lead times with an exponential delay, which generates second order systems independent of the lead times. If, on the other hand, we relax this assumption and model lead times through pure delays (as in the models presented in chapters 3 and 4), the order of the system depends on the lead time (Disney, 2008). This generates transcendental transfer functions that make a generalized analysis impossible. In his paper, Disney (2008) looks at the problem of rogue seasonality from a normative perspective; rather than describing the appearance of rogue seasonality, he takes the opposite approach and defines regions where the absence of the phenomena is guaranteed. He uses the 'Jury's inners' method to completely describe the regions for which a discrete system with a fixed lead time (with $L = 2$) exhibits an aperiodic response (with no oscillations) or a weakly-aperiodic response (with a limited number of oscillations). In Chapter 4, we show that for lead times larger than 2, the high-order nature of the system introduces oscillatory components for all but a subset of parametrizations (the so called DE-policies we studied in Chapter 4). However, this approach only characterizes the impulse response of the system, equivalent to a one-period spike in demand.

In the empirical literature, we find a series of efforts to quantify artificial seasonality (Lai, 2008) as well as the effect of demand seasonality in inventory variance (Steinker and Hoberg, 2012), and the effects of controlling for seasonality when measuring the Bullwhip Effect (Cachon et al., 2007; Gorman and Brannon, 2000). These empirical studies, due to the limitations of available data, can only analyze low frequency cycles: Quarterly cycles the case of Lai (2008) and Steinker and Hoberg (2012), and yearly cycles in the case of Cachon et al. (2007). Lai finds evidence of artificial seasonality being generated by the fiscal year effect (FYE) (where inventories tend to decrease in the last quarter of a fiscal year), however, this seasonality is thought to be due to specific executive action rather than due to a systematic effect stemming from an order-generation process. Steinker and Hoberg (2012) study the effect of numerous explanatory variables on the volatility of firm-level inventories and find evidence that this volatility increases with sales seasonality. They suggest that this fact supports the hypothesis that firms smooth production, using inventories to buffer against the seasonal changes in demand. Cachon et al. (2007) argue that because demand seasonality represents a strong incentive for firms to smooth production, all empirical studies should be conducted with demand series that are not corrected for seasonality. Pragmatically, they contend that "firms must produce to meet demand, not seasonally adjusted demand" and show that the Bullwhip Effect is over-estimated when using seasonally adjusted data. These studies, showing its immediate practical relevance, motivate further research to understand how seasonality affects inventory models and vice versa.

In this chapter, we perform extensive numerical experimentation on our discrete-time model to investigate the relationship between the behavioral parameters and the appearance –or attenuation– of rogue seasonality for different demand streams. In this view, in line with the thinking behind Chapter 4, we approach the problem in a descriptive manner: Not looking for optimal parameterizations, but rather to describe the interplay between behavior, demand patterns, and rogue seasonality. We adopt the methodology of Spectral Principal Component Analysis (SPCA) to achieve our objective. SPCA is an application of Principal Component Analysis (PCA) to spectral data. PCA is a widespread statistical tool that exploits orthogonal transformations to convert multi-dimensional data into useful representations in a lower dimensional space. In the operations research literature, SPCA was first used by Thornhill and Naim (2006) to demonstrate, through a case study, the use of SPCA to qualitatively identify rogue seasonality in a supply network. On a recent study, Shukla et al. (2012) develop a numerical method for rogue seasonality detection which they apply to a series of statistically generated demand streams, and test with real-life data from the case study of Thornhill and Naim (2006). They quantify, through their rogue seasonality index, the findings of the earlier study. We motivate our interest in rogue seasonality as an additional performance measure by showing that the popular metric used in the control systems literature (the Bullwhip measure, see Equation (4.24)) does not tell the complete story when the system observes actual demand realizations.

We carry out the numerical experimentation on discrete-time control-theoretic models and extend the SPCA methodology to quantify seasonality through a rogue seasonality index. For the design of experiments, because of the large two-dimensional parameter space, we fix the structural parameters and use concepts introduced in Chapter 4 to divide the parameter space into one-dimensional isometric Bullwhip Contours. We show that the findings hold for different parameter combinations in Appendix B.

The rest of this Chapter is structured as follows: In the next section we revisit the discrete-time control theoretic Automatic Pipeline Variable Inventory Order Based Production Control System (APVIOBPCS) model, methodologies for measuring the Bullwhip, and the SPCA methodology for rogue seasonality detection. In §5.3, we carry out the numerical experimentation and analysis of within-contour and between-contour behaviors. We present conclusions and insights derived from our work in §5.4. In an appendix, we present the performance of the system with varying structural parameters.

5.2 Model and methodology

In this section we introduce the discrete-time simulation model and the methodology which we use to use to quantify the cyclicality of the responses: Spectral Principal Component Analysis.

5.2.1 Discrete-time model

The numerical experiments in this chapter are based upon the discrete-time, periodic-review, single-echelon, general APVIOBPCS design that we introduced in Chapter 4. The inventory coverage ($C \in \mathbb{R}^+$), the delivery lead time ($L \in \mathbb{N}$), and the forecast smoothing parameter $\alpha \in [0,1]$ are the structural parameters of the system, while the pipeline ($\gamma_P \in [0,1]$) and inventory ($\gamma_I \in [0,1]$) adjustment factors are the behavioral parameters of the system. The inventory coverage represents the target inventory –measured in periods– that a firm chooses to maintain, while the target pipeline is calculated each period as the product of the forecast and the systems lead time. The lead time is assumed deterministic and defined as the time elapsed between the placement and receipt of a replenishment order. The behavioral parameters specify the fraction of the gap between target and actual values that are taken into account in the periodical ordering adjustment: γ_I is the fraction of the inventory gap to be closed per period, and γ_P is the fraction of the pipeline gap to be closed per period. For instance, a system with $\gamma_I = 1$ and $\gamma_P = 0$ completely closes the inventory gap every period, while it ignores the pipeline entirely.

Formally, the sequence of events and the equations in the model are as follows: at the beginning of each period (t) a replenishment order (o_t) based on the previous period's demand forecast (f_{t-1}) is placed with the supplier. Following this, the orders that were placed L periods before are received. Next, the demand for the period (d_t) is observed and served. Excess demand is back-ordered. Then, the demand forecast is updated according to the formula: $f_t = \alpha d_t + (1-\alpha) f_{t-1}$. The forecast is used to compute the target levels of both inventory, $\hat{i}_t = C f_t$, and pipeline, $\hat{p}_t = L f_t$. The orders that will be placed in the following period (o_{t+1}) are generated according to an anchor and adjustment-type procedure, $o_{t+1} = \gamma_I(\hat{i}_t - i_t) + \gamma_P(\hat{p}_t - p_t) + f_t$. The balance equations for inventory (i) and pipeline (p) are: $i_t = i_{t-1} + o_{t-l} - d_t$, and $p_t = p_{t-1} + o_t - o_{t-l}$.

In Chapter 4, we used control theory to analyze this model and developed insights regarding the influence of behavior on a set of performance measures. In this Chapter, we perform extensive numerical experiments on the model to (a) benchmark the performance measures against the actual behavior of the system, and (b) develop insights regarding a different dimension of the system's performance; rogue seasonality. Figure 5.1 shows a schematic of the model's block diagram. We refer the reader to Chapter 4 for the explicit derivation of the inventory and order transfer functions.

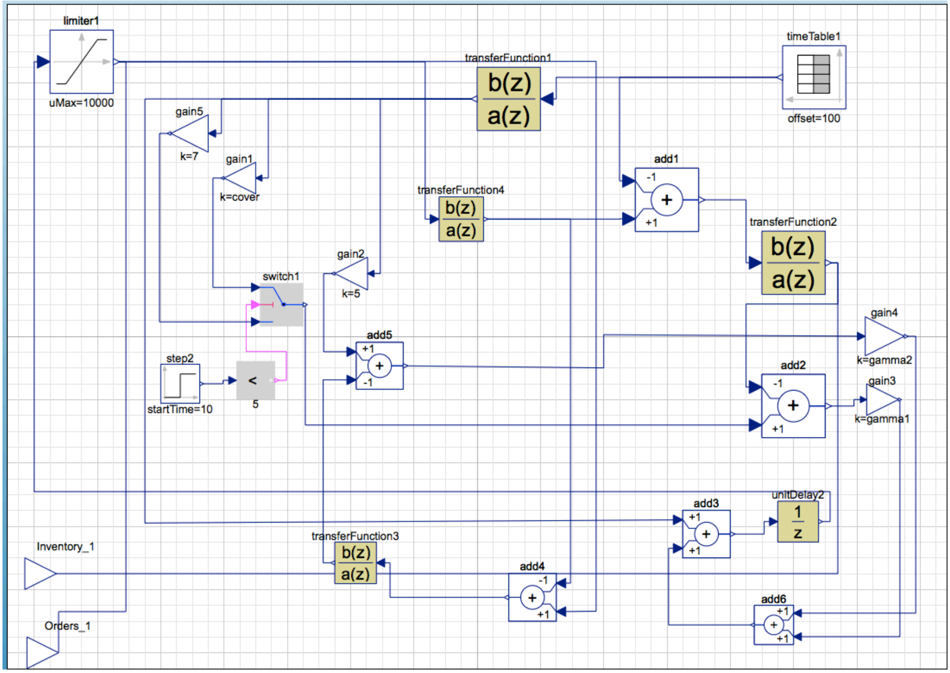


Figure 5.1 *Block diagram of model as implemented in software.*

5.2.2 *Spectral Principal Component Analysis*

The seasonal information of the response of a system is contained in its frequency plot. Systems with similar seasonal components have similar frequency plots; peaks at a given frequency reveal a seasonal component at that frequency. Visual analysis of the frequency response is a powerful tool to compare the cyclicality/dampening of the response of a system with respect to its input (or among different systems), however, this is a qualitative methodology.

Spectral Principal Components Analysis provides us with the methodology needed to quantify similarities and differences between the frequency response of different behavioral policies. Intuitively, SPCA reduces the dimensionality of a frequency response plot by quantifying a single dimension that captures its most prominent frequency characteristic. It maps an entire frequency spectrum to a point in a 2- or 3-D space. Thus, using SPCA we can characterize how the dynamic response of the system changes to changes in parameterizations and demand patterns through a graphical representation of these frequency plot mappings.

SPCA describes the application of Principal Component Analysis (PCA) to spectral data. The core idea behind PCA is to reduce the dimensionality of a data set consisting of a large number of variables, while retaining as much of the variation present in the dataset as possible. This is achieved by first transforming the data to a new set of uncorrelated variables (the Principal Components) that are ordered in such a way that a the first few retain most of the variation of the original data (Jolliffe, 2005). Since we can describe most of the variation in the original data with the first few variables, if we discard the remaining variables we effectively reduce the dimensionality of the problem without losing valuable information. In the case of our control theoretic inventory model, the dimensionality of the data is the amount of distinct frequencies that we use to compute the systems' frequency response. We use SPCA to condense all the information contained in a given frequency plot into a point on a 2D plane or 3D space.

Principal Component Analysis is used in a wide variety of fields; for example, Shlens (2005) presents a clear explanation of the method including examples from the world of physics, and Dougherty (2013) introduces it as a method for pattern recognition. SPCA was first described in a patent application by Belchamber et al. (1992) as a method to identify and cluster acoustic emissions through a measured spectrogram (essentially automating the detection of faulty motors by the sound they produce). Recently, Thornhill et al. (2002) use SPCA to cluster different chemical processes within a plant, and Thornhill and Naim (2006) use the methodology to cluster different components of a supply network according to the oscillatory components. We adopt their implementation of SPCA, which we detail below.

To compare the responses of different systems or parameterizations, we follow these steps:

1. Collect the time-series data of all the experiments (model runs).
2. Mean-center the time-series data and calculate each of their single-sided power spectra.
3. Scale the spectra so that the total power is unity.
4. Perform the PCA decomposition of the data.
5. Calculate contribution of each of the Principal Components to the total

variance.

6. Keep the limited number of PC's that explain most[1] of the variation.

Let x_i be the power spectra of experiment i, where the power, P, is measured at n different frequencies:

$$x_i = (P_i(f_1), \ldots, P_i(f_n)) \tag{5.1}$$

We can then, when we have m experiments, construct the $m \times n$ matrix X:

$$X = \begin{pmatrix} P_1(f_1) & \ldots & P_1(f_n) \\ \vdots & \ldots & \vdots \\ P_m(f_1) & \ldots & P_m(f_n) \end{pmatrix} \tag{5.2}$$

We can interpret this as the matrix that results in arranging one experiment per row and one frequency per column. The PCA decomposition transforms the matrix X into the sum over m orthonormal basis functions w'_1 to w'_m. These are row vectors with N frequency channels. Intuitively, one can think of each of the w'_i vectors as a series of weights placed on each of the original N frequencies. The column vectors $(t_{1,i}, \ldots, t_{m,i})'$ are the principal components:

$$X = \begin{pmatrix} t_{1,1} \\ \vdots \\ t_{m,1} \end{pmatrix} w'_1 + \begin{pmatrix} t_{1,2} \\ \vdots \\ t_{m,2} \end{pmatrix} w'_2 + \ldots + \begin{pmatrix} t_{1,m} \\ \vdots \\ t_{m,m} \end{pmatrix} w'_m \tag{5.3}$$

We can rewrite Equation (5.3) in a compact way as $X = TW'$ where the *ith* column of T is the vector $(t_{1,i}, \ldots, t_{m,i})'$ and the rows of W' are w'_1 to w'_m. Because, by definition, the rows W' are orthonormal, then W' is an orthogonal matrix and $W'W = I$. Thus, $T = XW$.

We use Singular Value Decomposition to compute the necessary vectors. Let $X = U\Sigma V'$ with $T = U\Sigma$ and $W = V'$. Knowing that matrix Σ is diagonal and its elements are the positive square roots of the eigenvalues of $X'X$. For experimental data, SVD is typically applied in software.

Equation (5.3) transforms the data without any loss of information. We can describe the majority of the variation of X by truncating the PCA description; if all the experiments share similar frequency traits, most of the variation can be described as a function of a limited amount of terms (principal components).

[1] This is a subjective assessment. As a reference, the articles cited above set the threshold at 80-90% of the variation

For example, we can rewrite X in terms of two principal components:

$$X = \begin{pmatrix} t_{1,1} \\ \vdots \\ t_{m,1} \end{pmatrix} w'_1 + \begin{pmatrix} t_{1,2} \\ \vdots \\ t_{m,2} \end{pmatrix} w'_2 + E, \tag{5.4}$$

Where E is an error matrix. In doing so, we reduce the problem of comparing n different frequencies of m different experiments to comparing just two principal components, $t_{i,1}$ and $t_{i,2}$ for each of the m experiments. This can be represented graphically by a 2D scatter plot in which each experiment maps to a point in the space defined by the two orthonormal basis w'_1 and w'_2. Therefore, by comparing the 2D representation of the principal components we quantify the relative similarity between the experiments' response, and by observing the form of the basis vectors w'_i we can gain insights and quantify the cyclical component of each experiment.

5.2.3 *Rogue seasonality detection*

Rogue seasonality describes "[cyclical] demand patterns that are induced by the internal processes themselves and not by any external influences" (Thornhill and Naim, 2006); to detect rogue seasonality in the output of a system we need then 2 things: a cyclical response, and for this cyclicality to be generated endogenously (i.e., not present in the input). Our approach is inspired on the observation of Shukla et al. (2012): When rogue seasonality is present, the different outputs of the system (inventory and orders) will become cyclical and similar to each other, while at the same time remain different from the input. We adopt their approach and calculate a rogue seasonality index based on the similarity of the responses among themselves and the input demand. Shukla et al. (2012) show that a rogue seasonality index constructed as

$$\frac{\text{Average dissimilarity between input and other variables}}{\text{Average dissimilarity between all variables}}$$

yields a useful and robust measure of the system's rogue seasonality. In their study, they measure dissimilarity as the euclidean distance between the variables in various transformed forms. We adopt the euclidean distance of the SPCA transformation as a measure of similarity between frequency responses. Formally, we define the rogue seasonality index for experiment i, ϕ_i, as:

$$\phi_i = \frac{\frac{1}{2}\left(\overline{d_{I_i-D}} + \overline{d_{O_i-D}}\right)}{\overline{d_{I_i-O_i}}}, \tag{5.5}$$

where $\overline{d_{I_i-D}}$ is the euclidean distance between the SPCA description of the inventory response of experiment i and the SPCA description of the demand; $\overline{d_{O_i-D}}$ is the euclidean distance between the SPCA description of the order response of experiment i and the SPCA description of the demand; and $\overline{d_{I_i-O_i}}$ is the euclidean distance between the SPCA description of the inventory response of experiment i and the SPCA description of the order response of experiment i . Additionally, to better understand the rogue seasonality of the system, we define two additional measures:

$$\phi_{I,i} = \frac{\overline{d_{I_i-D}}}{\overline{d_{I_i-O_i}}}, \tag{5.6}$$

$$\phi_{O,i} = \frac{\overline{d_{O_i-D}}}{\overline{d_{I_i-O_i}}}, \tag{5.7}$$

with $\phi_{I,i}$ and $\phi_{O,i}$, we measure the similarity of each of the responses individually. In doing this, we can detect whether either orders or inventories are independently exhibiting rogue seasonality in their response.

5.3 *Numerical Results*

This section is divided as follows: We first introduce an experimental design based on the behavioral parameters and the stationary response of the system; then, we present the demand streams that we use as an input for the experiments; we revisit the Bullwhip (stationary) performance measures derived Chapter 4 through a comparison of the theoretical and numerical results; finally, we study the dynamic performance of the systems by comparing the rogue seasonality of different behavior-demand combinations through the SPCA methodology.

5.3.1 *Experimental Design*

The system has 3 structural parameters (C, L, and α) and 2 behavioral parameters (γ_I and γ_P). In Chapter 4, we fixed the structural parameters to study the influence of the behavioral parameters on the stationary and dynamic performance measures. Following on this, in this section we fix the structural parameters and use the insights from the stationary performance of the system to develop an experimental design for the behavioral parameters. We present a set of performance metrics for various values of structural parameters in Appendix B; These metrics suggest that the insights obtained from the analysis of behavioral changes hold under different structural parameters.

Parametrization of experiments To create the experimental designs we fix the structural parameters ($C = 3, L = 5, \alpha = 0.3$) and calculate isometric-Bullwhip lines for the entire behavioral parameter space—based on BW_O, the theoretical stationary measure derived in Chapter 4. We then define 4 different experiment clusters, consisting each of 18 parameterizations along each of the contours. The 73 different parameter combinations that belong to the 4 different clusters are shown in Figure 5.2, and the values of their Bullwhip of orders, in Table 5.1.

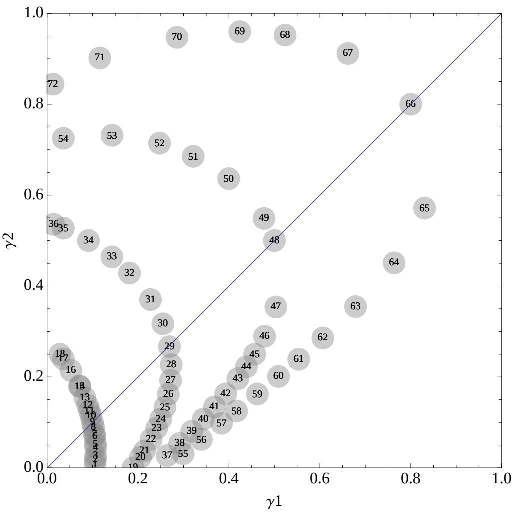


Figure 5.2 *Experiments in the behavioral parameter space.*

Table 5.1 *Experimental Design*

Cluster	Experiments	BW_O
1	1-18	0.7
2	19-36	2.0
3	37–54	4.6
4	55-72	9.7

Thus, all parameterizations that belong to a given cluster have the exact same BW_O for orders as calculated through Equation (4.24). Table 5.2 shows the behavioral parameters, the theoretical stationary performance measures (BW_O and BW_I) and the experimental dynamic measures ($ITAE_O$ and $ITAE_I$, see Equation (4.26)). The bold results are those that conform to the DE-diagonal, where $\gamma_I = \gamma_P$ and the experiments followed by an asterisk (*) correspond to experiments in which the dominant pole of the transfer function is real (i.e., experiments with a dampened response).

Table 5.2 *Design of experiments*

(a) *Experimental cluster 1*

	γ_I	γ_P	BW_o	BW_I	$ITAE_o$	$ITAE_I$
1	0.11	0	0.7	10.3	13.0	131.9
2	0.11	0.02	0.7	10.3	12.4	129.7
3	0.11	0.03	0.7	10.2	11.8	128.4
4	0.11	0.04	0.7	10.1	11.2	129.1
5	0.11	0.05	0.7	10.0	10.9	134.2
6	0.10	0.07	0.7	9.9	10.9	135.0
7	0.10	0.07	0.7	9.9	10.8	43.7
8	0.10	0.09	0.7	9.8	10.8	51.7
9	**0.10**	**0.1**	**0.7**	**9.7**	**10.7**	**162.1**
10*	0.10	0.11	0.7	9.6	10.7	178.5
11*	0.09	0.13	0.7	9.5	10.7	187.3
12*	0.09	0.14	0.7	9.4	10.6	204.4
13*	0.08	0.15	0.7	9.3	10.5	260.5
14*	0.07	0.18	0.7	9.1	10.5	264.7
15*	0.07	0.18	0.7	9.1	10.2	325.9
16*	0.05	0.21	0.7	8.9	9.6	410.6
17*	0.04	0.24	0.7	8.7	8.4	529.0
18*	0.03	0.25	0.7	8.7	2.3	927.0

(b) *Experimental cluster 2*

	γ_I	γ_P	BW_o	BW_I	$ITAE_o$	$ITAE_I$
19	0.19	0.00	2.0	14.7	28.4	160.3
20	0.21	0.02	2.0	14.8	24.9	141.3
21	0.21	0.04	2.0	14.8	23.0	132.4
22	0.23	0.06	2.0	14.7	20.2	119.0
23	0.24	0.09	2.0	14.6	18.8	112.4
24	0.25	0.11	2.0	14.5	15.9	100.3
25	0.26	0.13	2.0	14.2	14.8	96.2
26	0.27	0.16	2.0	14.0	13.1	89.9
27	0.27	0.19	2.0	13.6	12.0	86.4
28	0.27	0.23	2.0	13.2	10.8	83.5
29	**0.27**	**0.27**	**2.0**	**12.8**	**10.0**	**84.9**
30*	0.25	0.32	2.0	12.2	10.0	90.6
31*	0.23	0.37	2.0	11.6	10.0	104.8
32*	0.18	0.43	2.0	11.1	10.0	124.7
33*	0.14	0.46	2.0	10.9	10.0	167.9
34*	0.09	0.50	2.0	10.6	10.0	220.0
35*	0.04	0.53	2.0	10.5	10.1	309.1
36*	0.01	0.54	2.0	10.4	9.4	479.7

(c) *Experimental cluster 3*

	γ_I	γ_P	BW_o	BW_I	$ITAE_o$	$ITAE_I$
37	0.26	0.03	4.6	18.9	55.9	248.3
38	0.29	0.05	4.6	19.4	49.9	216.6
39	0.32	0.08	4.6	19.7	43.5	185.8
40	0.34	0.11	4.6	19.9	38.4	162.7
41	0.37	0.13	4.6	20.0	33.8	143.3
42	0.39	0.16	4.6	20.0	29.5	125.3
43	0.42	0.20	4.6	19.8	26.1	112.2
44	0.44	0.22	4.6	19.6	23.0	101.8
45	0.46	0.25	4.6	19.3	20.5	92.8
46	0.48	0.29	4.6	18.8	17.6	83.4
47	0.50	0.35	4.6	17.8	13.8	72.5
48	**0.50**	**0.50**	**4.6**	**15.5**	**10.0**	**67.3**
49*	0.48	0.55	4.6	14.8	10.0	71.3
50*	0.40	0.64	4.6	13.8	10.0	82.8
51*	0.32	0.69	4.6	13.4	10.0	104.4
52*	0.25	0.71	4.6	13.3	10.2	148.8
53*	0.14	0.73	4.6	13.3	11.1	222.4
54*	0.04	0.73	4.6	13.4	12.3	467.9

(d) *Experimental cluster 4*

	γ_I	γ_P	BW_o	BW_I	$ITAE_o$	$ITAE_I$
55	0.30	0.03	9.7	22.2	88.2	352.4
56	0.34	0.06	9.7	23.4	81.8	311.2
57	0.38	0.10	9.7	24.6	76.7	280.2
58	0.42	0.13	9.7	25.3	69.4	246.0
59	0.46	0.16	9.7	26.2	63.5	219.2
60	0.51	0.20	9.7	26.8	56.1	189.8
61	0.55	0.24	9.7	27.3	50.6	168.8
62	0.61	0.29	9.7	27.5	44.8	149.4
63	0.68	0.35	9.7	27.2	40.3	133.7
64	0.76	0.45	9.7	25.8	29.9	102.9
65	0.83	0.57	9.7	23.0	20.8	77.7
66	**0.80**	**0.80**	**9.7**	**18.3**	**11.7**	**58.9**
67*	0.66	0.91	9.7	17.0	10.0	62.8
68*	0.52	0.95	9.7	16.7	10.0	98.7
69*	0.42	0.96	9.7	16.9	10.4	121.9
70*	0.29	0.95	9.7	17.4	11.7	165.5
71*	0.12	0.90	9.7	18.8	16.8	353.7
72*	0.01	0.84	9.7	18.9	20.0	669.7

5.3.2 *Response Types*

We saw in Chapter 4 that imaginary poles generate oscillatory responses and positive and real poles, dampening. In the proof of Proposition xxx we characterized the structure of the poles in the following way,

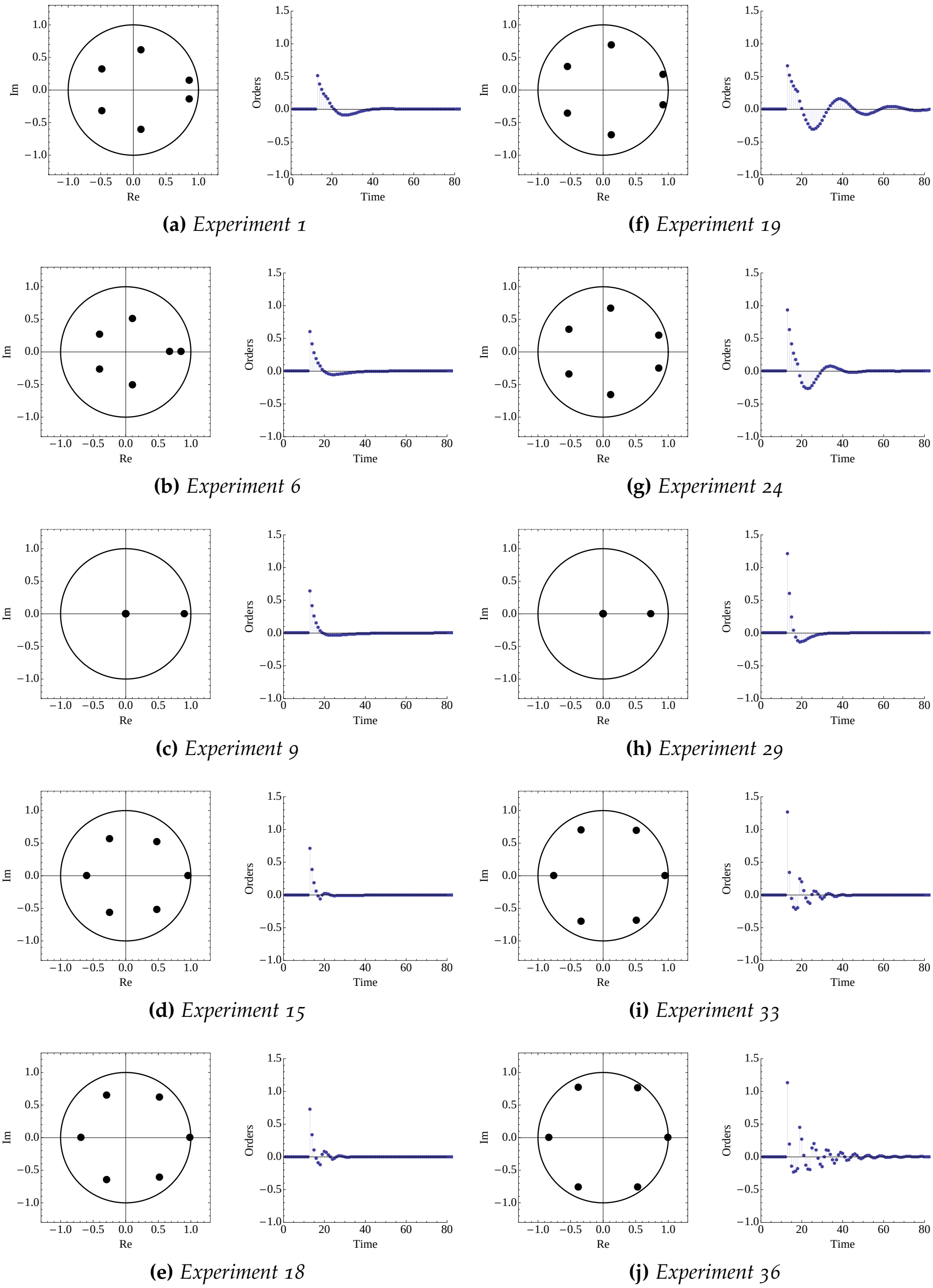
Definition 5.1 *When $\gamma_I > \gamma_P$ there is always a real and positive pole.*
When $\gamma_I < \gamma_P$ there is either zero or two a real and positive poles.

When a system is of order larger than 1, the response of the system , however we can approximate the response by analyzing the dominant pole—the pole with the largest real part. When the dominant pole is real, the system will exhibit a dampened response. When the dominant poles are a pair of complex conjugates, the system will exhibit an oscillatory response (Nise, 2007). To build up intuition with regards to the time-domain response of our system with different behavioral parameters, we can plot the poles resulting from different experimental designs in the complex plane. Detecting the dominant pole of each is thus immediate. Additionally, we can plot the time-domain response to an impulse input function.

In Figure 5.3, spread over pages 116 and 117, we plot the poles and impulse response of a number of experiments. We group clusters in four columns with two plots each, the leftmost column shows experiments from Cluster 1 and the rightmost column, experiments from Cluster 4. Poles (left of each column) are plotted on the complex plane. The time response (right of each column) corresponds to an input of a unit impulse. The top two rows correspond to $\gamma_I > \gamma_P$, the middle row corresponds to $\gamma_i = \gamma_P$, and the two bottom rows correspond to $\gamma_I < \gamma_P$. Poles (left) are plotted on the complex plane. The time response (right) corresponds to an input of a unit demand impulse. We clearly see the influence of the location of the dominant pole in the time response.

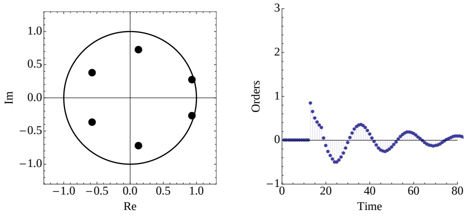
When the dominant poles are a pair of complex conjugates (figures 5.3a, 5.3f, 5.3g, 5.3k, 5.3l, 5.3p, and 5.3q), the response is clearly oscillatory. Note that the magnitude of the imaginary component of the dominant pole pairs increases towards the right, and that this increase is accompanied by an increase in the amplification and persistence of the oscillations[2]. When all poles are real (figures 5.3c, 5.3h, 5.3m, and 5.3r), the response is aperiodic. Finally when, in the presence of complex poles, the dominant pole is real, then the response is of an over-dampened oscillation (figures 5.3d, 5.3e, 5.3i, 5.3j, 5.3n, 5.3o, 5.3s, and 5.3t).

[2]Note that the scale of the time plots of the two rightmost columns is double that of the two leftmost columns, to account for the increased amplification.

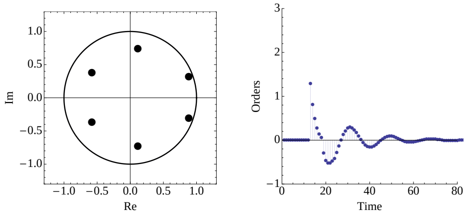


(a) *Experiment 1*
(f) *Experiment 19*
(b) *Experiment 6*
(g) *Experiment 24*
(c) *Experiment 9*
(h) *Experiment 29*
(d) *Experiment 15*
(i) *Experiment 33*
(e) *Experiment 18*
(j) *Experiment 36*

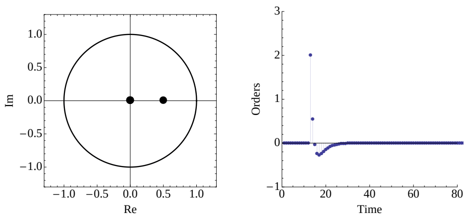
Figure 5.3 *Poles and time response of selected experiments.*



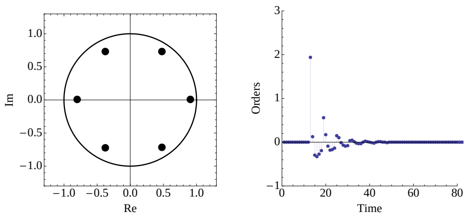
(k) *Experiment 37*



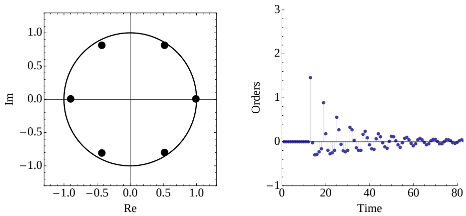
(l) *Experiment 42*



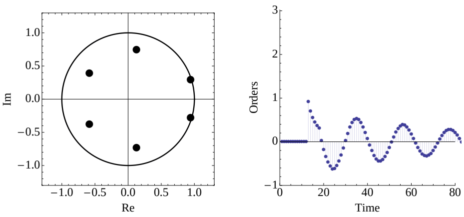
(m) *Experiment 48*



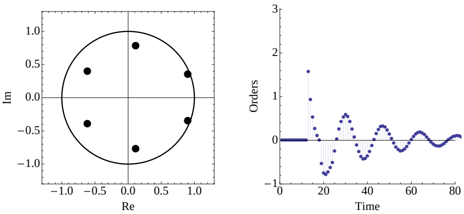
(n) *Experiment 51*



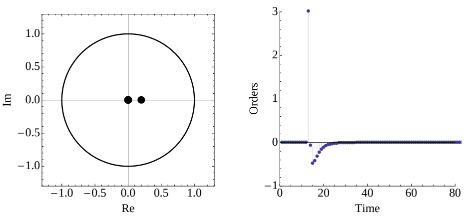
(o) *Experiment 54*



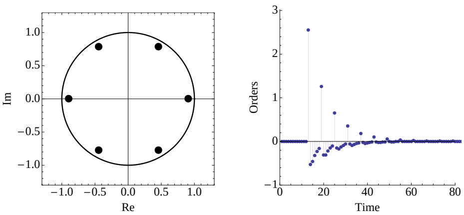
(p) *Experiment 55*



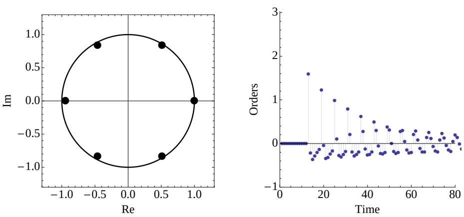
(q) *Experiment 60*



(r) *Experiment 66*



(s) *Experiment 69*



(t) *Experiment 72*

Figure 5.3 *Poles and time response of selected experiments.*

5.3.3 System Response Under Different Demand Conditions

The performance measures shown in Table 5.2 make strong assumptions regarding the underlying demand stream. The Bullwhip measures (BW_O and BW_I) assume a normal demand and the transient measures ($ITAE_O$ and $ITAE_I$), a single unit-step change in demand. It is unlikely that an production/inventory system encounters in real-life a demand with one of these characteristics. In chapter 4, we studied the trade-off of stationary and transient measures independently for orders and inventories. We found that, picturing the whole parameter space in two dimensions, both BW_O and BW_I decrease when the behavioral parameters approach the lower left quadrant, and increase when they approach the upper right quadrants. The opposite is true of the transient measures, $ITAE_O$ and $ITAE_I$. Additionally, we found that systems with behavioral parameters along the DE-diagonal offer a reasonable performance trade-off, systems below the DE-diagonal offer marginally better performance in certain instances with the downside of increased sensitivity to parameters, and systems above the DE-diagonal offer robustness to parameters with the cost of inferior performance (Figure 4.6).

The thinking behind the study of the trade-off between the stationary and transient measures is that whereas real-life demand does not consist of pure white noise nor of pure shocks, systems that perform well under both extreme conditions will perform well in real life. To study whether this assessment is adequate, we perform numerical experiments and measure the performance of the system under varying behavioral conditions (as per our experimental clusters) and demand streams.

5.3.4 Robustness of the Bullwhip Measure

We test the robustness of the theoretical bullwhip measures by running a series of simulations using a demand sampled from a normal distribution as an input. Table 5.3 shows the measured bullwhip for orders and inventory ($\overline{BW_O}$, $\overline{BW_I}$) for each of the experimental clusters and the respective percent error when compared to the theoretical measures ($\Delta_O\% = (\overline{BW_O} - BW_O)/BW_O$ and $\Delta_I\% = (\overline{BW_I} - BW_I)/BW_I$). The input to the simulation in each of the experiments is a demand stream consisting of 10000 observations sampled from a normal distribution with $\mu = 0$ and $\sigma = 1$[3].

The average absolute percent error of orders ($|\Delta_O\%|$) is 3.9% and its median

[3]The theoretical bullwhip measure is insensitive to the demand parameters, its only requirement is the input to consist of "white noise", that is, for all frequencies to be equally represented—a trait of the normal distribution.

2.2%. In the case of $|\Delta_I\%|$, the discrepancy between theoretical and actual values is larger, with an average of 50.8% and a median of 22.6%. If we look at the extreme values, we see that the minimum difference in both cases is 0%, but the maximum deviations are 12.4% for orders and 462.9% for inventories. The large spread of the error in inventories suggests that they are more sensitive than orders to changes in behavioral parameters, and casts doubts over the usage of the theoretical bullwhip as a performance measure[4]. To control for the experimental design affecting the outcomes, we repeated the experiments with demand streams consisting of $100, 10^3, 10^5$, and 10^6 data points, as well as demands sampled using different random number seeds.

Every instance generates comparable results. This suggests that the difference in the frequency responses between the theoretical and sampled demands is significant enough that it introduces errors in the estimation of the bullwhip measure. Note that systematically, the discrepancy between theoretical and measured bullwhip is larger at the extremes of the experimental clusters; this corresponds to the observation experiments along the DE-diagonal exhibit the flattest frequency response for both orders and inventories, while extreme values introduce pronounced peaks at specific frequencies. Additionally, the frequency response of inventories for a given parameterization contains larger peaks than its corresponding frequency response of orders (see Appendix B). This explains the difference in magnitude of the inventory and order absolute percent errors.

Steady state performance revisited In §4.4.1, we used frequency response plots to illustrate how the behavioral parameters affect the performance of the system by amplifying different frequencies. For orders, we identified the DE-diagonal ($\gamma_I = \gamma_P$) as the flattest response, and showed that policies below the DE-diagonal ($\gamma_I > \gamma_P$) amplify the lower frequencies, while policies above the diagonal ($\gamma_I < \gamma_P$) exhibit peaks at higher frequencies. For inventories, on the other hand, we have seen that all policies amplify the lower frequencies and that policies above the DE-diagonal show an additional peak at higher frequencies.

The effect of the uneven amplification of the different frequencies of the input signal is such that the noise in demands sampled from theoretical distributions is enough to introduce substantial differences from the theoretical performance. This is consistent with the results we see in Table 5.3; the relatively flat response of policies near the DE-diagonal result in small deviations from the theoretical performance, while the frequency-specific peaks of policies close to the edge of the parameter space seem to drive the deviations. Inspection of Figures 4.3

[4]Remember that we are not yet measuring performance per se; solely the performance of the theoretical performance measures.

and 4.4 shows that: (a) for comparable policies frequency peaks in the inventory response tend to be larger than for the order response; (b) that this difference is greater when policies are closer to the lower-left quadrant of the parameter space; and (c) that the order response appears to be more sensitive to the variation of behavioral parameters.

Even though these characteristics of the frequency response help us understand several of the observations we have made, they are inherently qualitative and hard to interpret if the demand is anything but a pure sine wave or normally distributed—analyzing the frequency response of each of the possible policies is a useful, but impractical technique if we are interested in more general insights.

To overcome this weaknesses of frequency plot analysis, we use SPCA to decompose the frequency information into a handful of principal components that we can plot. With these plots, we learn about the way that behavioral parameters affect the seasonal component of the system. We further quantify the response through the rogue seasonality index ϕ, which condenses the graphical information into one dimension.

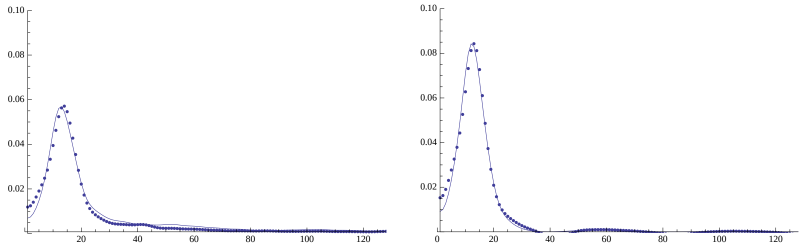
SPCA as a quantification of rogue seasonality We saw in §5.2.2 that by performing principal component analysis on spectral data, we can reduce its dimensionality with little loss of information. When we apply SPCA to the output of our system fed with the sampled normally distributed demand, we can describe 90.4% of the variance information using 3 principal components. By reducing the dimensionality of spectral data to 3 dimensions, we transform the subjective problem of comparing the frequency response plots of different policies into a quantitative one because we can represent each frequency plot by a single point in the 3-dimensional space. Figures 5.5 to 5.7 illustrate the application of SPCA to the output of our system. Instead of drawing cumbersome 3D plots, we represent the three dimensional space in the 'flatland' of the page by plotting three orthogonal projections called score plots. Each axis (score) represents one of the w' basis functions (as a convention, score 1 is the dimension that describes most of the variation in the data). Each point in the plot represents one spectrum in the data set, we use pink for the inventory spectra and blue for the order spectra. Additionally, we include a green point in the plots to represent the spectra of the input demand. Figure 5.5 plots score 1 Vs. score 2, Figure 5.6 plots score 2 Vs. score 3, and 5.7 plots score 1 Vs. score 3.

Each of the figures contains 4 individual plots: The top plot **(a)** aggregates the spectra of all 4 experimental clusters and each of the 4 bottom plots **(b)-(e)** represents a single experimental cluster. To distinguish between individual experiments, we identify the even-numbered experiments in the bottom graphs.

For each point in the plots, the coordinates represent the weights of t (as described in Equation (5.3)). For example, the coordinates for the order response of experiment 24 are $t_{24,1} \cong -0.085$, $t_{24,2} \cong 0$, and $t_{24,3} \cong 0.01$. Thus, multiplying each of the weightings by the w' vectors approximately reconstruct the spectrum. Figure 5.4 shows the spectrum of the order and inventory responses of experiment 24 and its reconstruction via the 3 principal components. Because a spectral plot represents the decomposition of a time-series into the sum of individual sine waves at different frequencies, similar spectral plots represent similar cyclical components. Thus, we know that the closer together two experiments are in the score plots, the more similar their seasonal component is: This is the basis of our rogue seasonality index measure.

The intuition is as follows. The green dot in each plot represents the input (demand) seasonality, each of the blue dots represents the order seasonality for a given experiment, and each of the pink dots represents the inventory seasonality for a given experiment. The rogue seasonality index measures the similarity in the response of orders and inventories among themselves and relative to the input demand. The rogue seasonality index increases the more dissimilar the frequency responses are among themselves. Thus, rogue seasonality decreases (a) the closer the blue, and pink, dots of a given experiment are to the green dot, and (b) the closer the blue and pink dots of a given experiment are among themselves.

In Figures 5.5 to 5.7 we see that the frequency response of the system with the different behavioral policies follows a smooth evolution within each contour (experimental cluster). This confirms that partitioning the parameter space in isometric contour lines is an efficient strategy. Also, we see that while the *type* of response variation across contours is similar, the *spread*, the difference between the extremes of each contour, appears to increase when for each successive experimental cluster. From these representations, we can glean several insights: Below the DE-diagonal (experiments $1 - 9$, $19 - 28$, $37 - 47$, and $55 - 65$) the response of orders and inventories is similar and distant from the input response (green dot). With each successive experiment, orders and inventories start to diverge; orders then show a frequency response that is closer to the that of the demand than it is to that of the inventory. This suggests that experiments below the diagonal exhibit a greater rogue seasonality (orders and inventory having similar frequency response, but different from the input), and that this rogue seasonality decreases when we move through a contour line. It is interesting to see how the frequency response of inventories does not approximate the frequency response of the demand—consequently remaining cyclical in all



(a) *Orders of experiment 24* **(b)** *Inventory of experiment 24*

Figure 5.4 *Reconstructed spectra (solid line) and complete spectra (dots)*

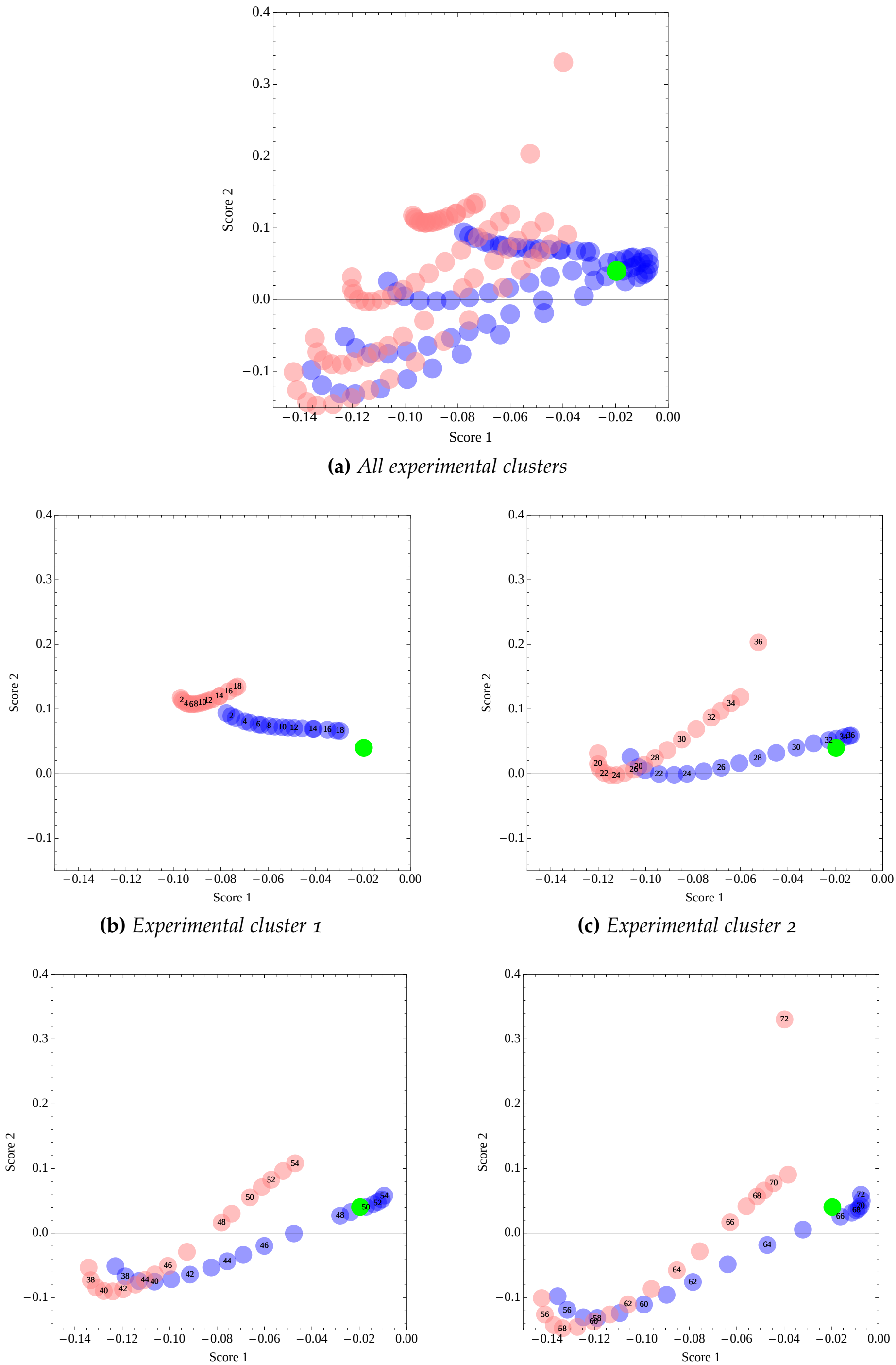
experiments.

In Table 5.4 we see the complete performance metrics of all experiments, including the seasonality indices. We can see that ϕ correctly captures the behavior we observe in Figures 5.5 to 5.7: Experiments below the DE-diagonal have a high rogue seasonality, this value decreases along each experimental cluster up to a minimum point, and then increases again. We also see that for a vast majority of the behavioral policies, inventories are more susceptible to rogue seasonality than orders. However, it is apparent that the measure of rogue seasonality in itself is not an absolute performance metric: As expected, a high seasonality index ($\phi > 1$) is followed by poor performance according to the bullwhip and ITAE metrics; however, when $\phi < 1$ (and thus the difference between the frequency response of orders and inventories is greater than the difference between these and the demand) the system also under-performs.

5.3.5 System performance with seasonal demands

Thus far, we have explored different aspects of the system's performance when demand is sampled from a normal distribution. We have found that the system's response depends on the input frequency: thus, even when the demand is sampled from a theoretical distribution, the different amplification peaks cause the actual performance to deviate from the theoretical measures. Therefore, the performance of the system will be intrinsically linked to the seasonality of the input demand. To be able to form a more complete picture of the system's response, we generate 3 additional demand streams by adding different artificial seasonal components to our sampled normally distributed demand.

We generate three artificial seasonal demands by adding the original sampled demand stream to three different sine waves. Formally, let $D_1(t)$ be the sampled



(a) *All experimental clusters*

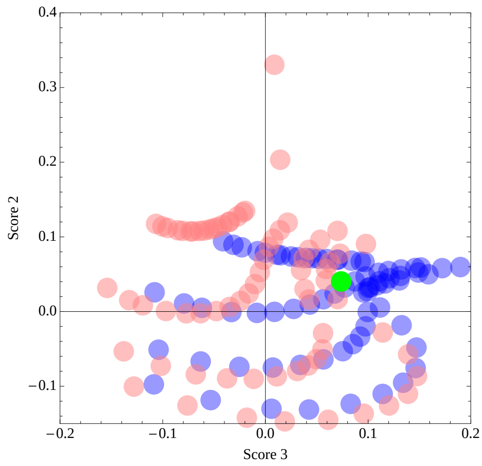
(b) *Experimental cluster 1*

(c) *Experimental cluster 2*

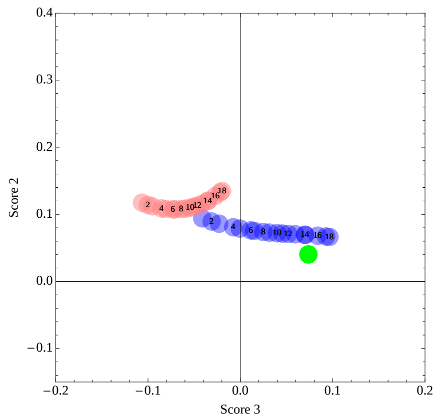
(d) *Experimental cluster 3*

(e) *Experimental cluster 4*

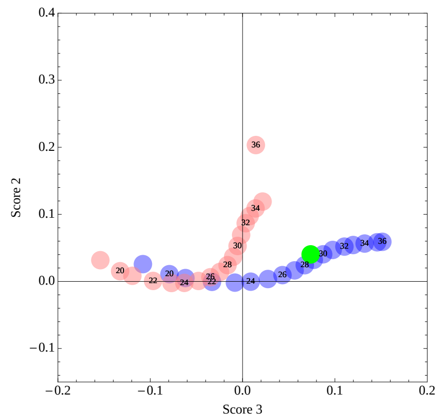
Figure 5.5 *Score plots of the SPCA decomposition of the frequency response of orders (blue) and inveventories (pink). First projection (Score 1 Vs. Score 2).*



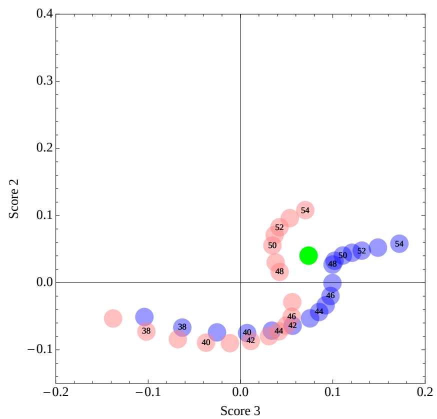
(a) *All experimental clusters*



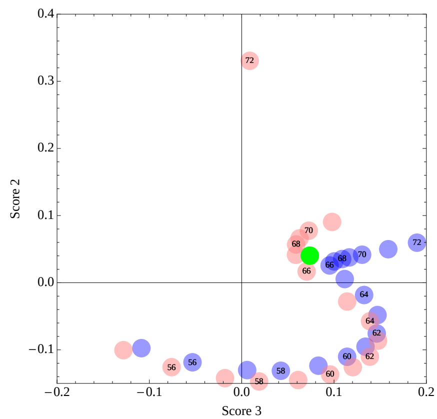
(b) *Experimental cluster 1*



(c) *Experimental cluster 2*

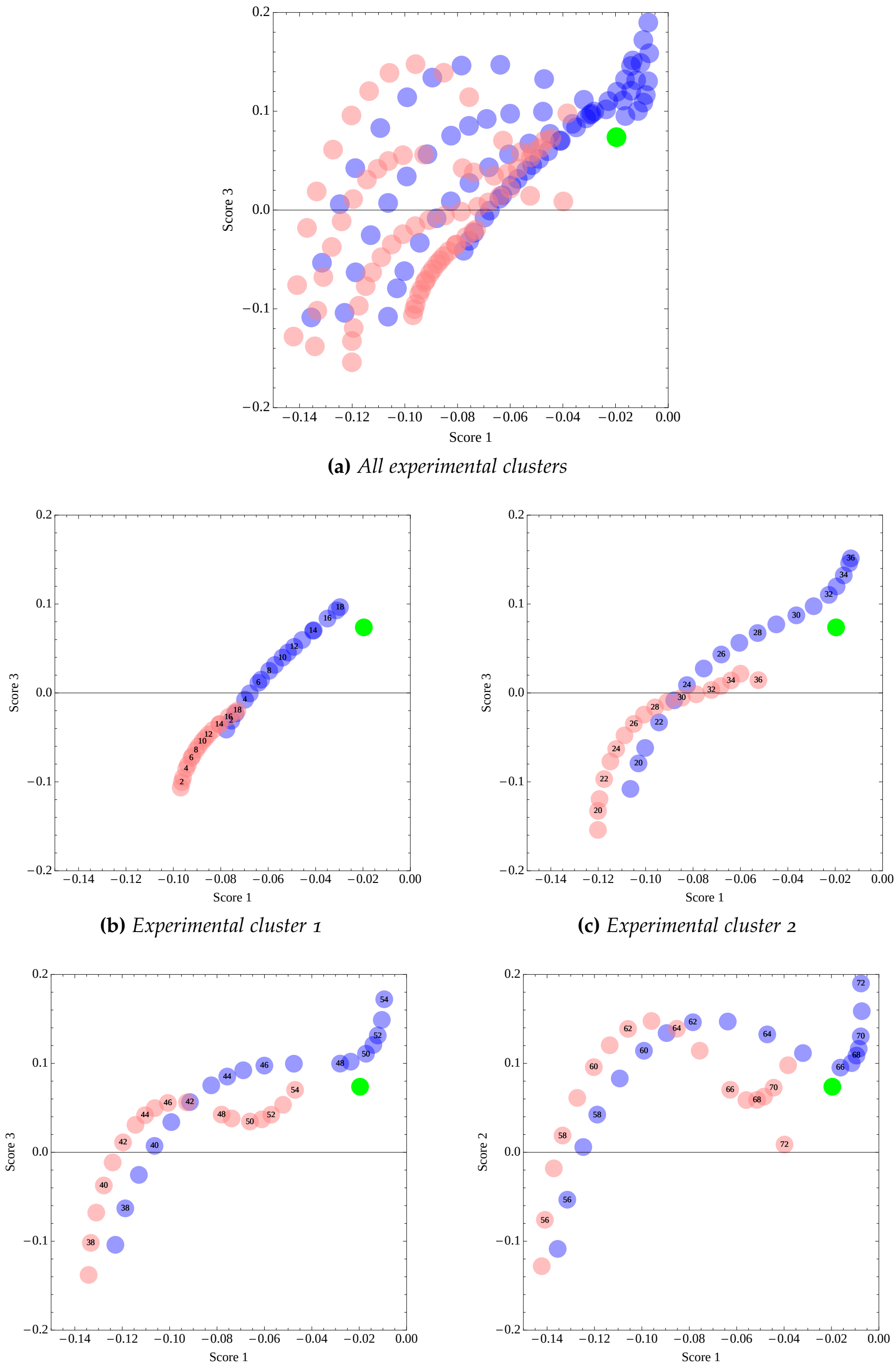


(d) *Experimental cluster 3*



(e) *Experimental cluster 4*

Figure 5.6 *Score plots of the SPCA decomposition of the frequency response of orders (blue) and inveevntories (pink). Second projection (Score 2 Vs. Score 3).*



(a) *All experimental clusters*

(b) *Experimental cluster 1*

(c) *Experimental cluster 2*

(d) *Experimental cluster 3*

(e) *Experimental cluster 4*

Figure 5.7 *Score plots of the SPCA decomposition of the frequency response of orders (blue) and inventories (pink). Third projection (Score 1 Vs. Score 3).*

demand at time t, then:

$$D_2(t) = \sin\left(\frac{2\pi}{7}t\right) + D_1(t) \quad t \in \mathbb{I}\,[0, 10000], \tag{5.8}$$

$$D_3(t) = \sin\left(\frac{2\pi}{30}t\right) + D_1(t) \quad t \in \mathbb{I}\,[0, 10000], \tag{5.9}$$

$$D_4(t) = \sin\left(\frac{2\pi}{90}t\right) + D_1(t) \quad t \in \mathbb{I}\,[0, 10000]. \tag{5.10}$$

This represents (for a system that observes demand each day) weekly, monthly, and quarterly seasonal demand streams (D_2, D_3, and D_4 respectively). Observing Figures 4.3 and 4.4, we see that the both the inventory and order plots show a peak at a frequency of approximately $\frac{\pi}{8}$. However, in the case of orders, the peak is larger for parameter combinations below the DE-diagonal. This, translated into cyclical terms means that the system exhibits peak amplification with a ciclicality of around 25 days (assuming demand is measured daily). Knowing this, we expect that when the system faces a demand with a similar frequency (such as demand D_3), the interaction between exogenous and endogenous seasonality components will result in poor performance. In contrast, when the demand is of significantly higher frequency (such as demand D_2), the system will attenuate its amplification and rogue seasonality. In the case of exogenous demand D_4, of a lower frequency, the amplification of low frequencies results in poor performance compared to the normally distributed benchmark, but better performance than when the exogenous and endogenous seasonalities are of approximately the same frequency.

These results show that the behavior of the system plays a central role in the performance of the system, not only by defining the stability and bullwhip measures, but also by characterizing the rogue seasonality of the system and the interaction with seasonal demands. From a managerial perspective, this result highlights the importance of understanding both the endogenous seasonality of one's own system, as well as any seasonality present in the demand stream. From a theoretical perspective, the fact that exogenous seasonality affects the performance of an inventory policy underscores the opportunity to include such factors in inventory models. Empirically, this is in line with the reasoning from Cachon et al. (2007), where the authors argue that since real-life firms react to demands that can very well be seasonal, the models used should take this into consideration.

5.4 Conclusions

In this study, we have performed extensive numerical experiments, using a control theoretic model, with the objective of benchmarking the performance of different behavioral policies reacting to different demand patterns. We motivated this study with the findings of Chapter 4, where we analyze an equivalent system and find defined performance trade-offs that depend on the demand. Because the stationary performance metric assumes a normal demand, we first test the robustness of this measure by simulating the response of the system when the demand is sampled from a normal distribution. According to our findings, inventories are especially sensitive to demand patterns; the maximum difference between the theoretical and actual inventory performance is $\sim 450\%$.
To understand the reasons for this discrepancy, we study the frequency response plot of the inventory response of this system; we find that lower frequencies (centered around $\frac{\pi}{8}$, or a cycle of approximately 25 periods) are consistently amplified, and that this amplification depends on the behavioral parameters. This uneven amplification of different frequencies has two immediate effects: The aforementioned discrepancy between theoretical and simulated performance, and the appearance of rogue seasonality—the generation of a cyclical output from a random input.

Thus, in the presence of a given demand, the cyclicality of the production/inventory system itself can potentially dominate over any seasonality present in the demand. To better understand this, we introduce Spectral Principal Component Analysis as a way to quantify the frequency response and therefore be able better describe the appearance of rogue seasonality for different experiments. We develop a rogue seasonality index, based on the SPCA decomposition, and find that it captures the way in which different policies are susceptible to cyclical amplification. In particular, we observe intense rogue seasonality when the pipeline is underestimated (policies below the 'optimal' DE-diagonal) and the opposite effect when the pipeline adjustment is taken into account more than the inventory adjustment (policies above the DE-diagonal). The consequence of these observations is immediately observed when we study the interaction between the demand's seasonality and the system's rogue seasonality. When the seasonality of demand is of a higher frequency than the system, then taking the pipeline into account more than the inventory adjustment decreases the bullwhip of the system, while the opposite is true when the demand is of a lower frequency than the system.

This study is of a descriptive nature; rather than optimizing the system, we observe its performance when different parameters change. With this in mind, we have succeeded in uncovering, via the rogue seasonality measure, another dimension of inventory performance. Our study shows that in addition to having an influence in the stationary and transient theoretical performance metrics, behavioral parameters affect the performance of the system in other, non trivial ways. Behavioral parameters fundamentally change the sensitivity of the system to the demand stream. Furthermore, when the demand contains a seasonal component, the behavioral parameters also determine the periodic response of the system: different parameter combinations attenuate and amplify different frequencies.

From an operational perspective, this implies that characterizing and tracking the seasonality of customer demand is of importance inasmuch as it defines the core response of the system. In addition to understanding the seasonal component of demand, our study calls for an understanding of the rogue seasonality of one's system. The rogue seasonality index contains information that is not captured by the often used stationary and transient performance metrics.
At a tactical level, understanding rogue seasonality allows managers to better understand the medium- to long-term evolution of inventories and orders, potentially affecting the way internal performance metrics work. In this view, the baseline from which to measure inventory performance is not stationary, but cyclical.

More research is needed to further understand the trade-offs between all the different performance metrics we have presented. There is a need, in particular, to introduce explicit cost and service considerations into control theoretic models. *How do the trade-offs presented in chapters 4 and 5 relate to cost considerations, and service level requirements? Is rogue seasonality inherently negative? How does an acceptable cost-driven trade-off differ from a performance-driven trade-off?*.

To understand the generality of our numerical results, we present an Appendix with frequency plots and bullwhip contours for various structural parameter combinations. From these, we see that of the structural changes, the exponential smoothing constant α and the inventory coverage C affect the performance of the system in magnitude but not in its structure. All else equal, the bullwhip of a system increases with α and C, as does the amplification of a particular demand frequency. In contrast, a change in the lead time L brings about a fundamental change in the response. The amount of peaks observed in the frequency plots, as well as their frequencies, depend on the lead time.

Table 5.3 *Realizations of the bullwhip measures*

(a) *Experimental cluster 1*

	$\overline{BW_o}$	$\overline{BW_I}$	$\Delta_O\%$	$\Delta_I\%$
1	0.68	12.9	-2.3	25.1
2	0.68	12.4	-2.3	20.7
3	0.68	12.0	-2.3	17.4
4	0.68	11.3	-2.2	11.5
5	0.68	10.9	-2.2	8.9
6	0.68	10.4	-2.1	4.8
7	0.68	10.2	-2.1	3.6
8	0.68	9.8	-2.0	0.2
9	**0.68**	**9.5**	**-1.9**	**-2.0**
10*	0.68	9.1	-1.7	-4.7
11*	0.68	8.9	-1.6	-6.6
12*	0.68	8.6	-1.5	-8.7
13*	0.68	8.3	-1.2	-11.2
14*	0.69	7.8	-0.9	-14.5
15*	0.69	7.8	-0.9	-14.7
16*	0.69	7.2	-0.4	-18.8
17*	0.69	6.9	-0.1	-21.5
18*	0.69	6.7	0.1	-22.6

(b) *Experimental cluster 2*

	$\overline{BW_o}$	$\overline{BW_I}$	$\Delta_O\%$	$\Delta_I\%$
19	1.88	27.88	-4.8	90.4
20	1.89	25.31	-5.2	71.3
21	1.89	23.85	-5.5	61.3
22	1.88	21.59	-5.8	46.5
23	1.88	19.75	-5.9	35.1
24	1.88	18.47	-5.8	27.6
25	1.89	17.01	-5.6	19.5
26	1.89	15.66	-5.2	12.3
27	1.90	14.4	-4.7	5.9
28	1.92	13.21	-4.0	0.0
29	**1.93**	**12.03**	**-3.2**	**-5.6**
30*	1.95	10.78	-2.2	-11.5
31*	1.98	9.69	-1.1	-16.7
32*	2.00	8.69	0.0	-21.8
33*	2.01	8.15	0.5	-24.9
34*	2.02	7.62	0.9	-28.2
35*	2.02	7.2	1.1	-31.1
36*	2.02	10.63	1.2	2.2

(c) *Experimental cluster 3*

	$\overline{BW_o}$	$\overline{BW_I}$	$\Delta_O\%$	$\Delta_I\%$
37	4.27	51.0	-6.2	170.1
38	4.21	45.2	-7.8	133.4
39	4.18	40.6	-8.7	106.2
40	4.17	37.0	-9.0	85.8
41	4.17	34.0	-8.9	70.0
42	4.19	31.3	-8.6	56.5
43	4.21	28.5	-8.1	44.1
44	4.23	26.7	-7.7	36.0
45	4.26	24.9	-7.2	28.9
46	4.29	22.6	-6.4	20.3
47	4.34	19.5	-5.1	9.4
48	**4.49**	**14.5**	**-2.2**	**-6.6**
49*	4.52	13.2	-1.4	-10.7
50*	4.58	11.4	-0.2	-18.0
51*	4.60	10.4	0.3	-22.6
52*	4.61	9.8	0.4	-26.2
53*	4.60	9.2	0.5	-30.3
54*	4.60	8.9	0.6	-33.1

(d) *Experimental cluster 4*

	$\overline{BW_o}$	$\overline{BW_I}$	$\Delta_O\%$	$\Delta_I\%$
55	8.98	101.1	-7.1	355.9
56	8.61	86.0	-11.4	267.5
57	8.52	75.5	-12.4	207.0
58	8.56	70.0	-11.5	176.7
59	8.54	62.8	-11.8	140.0
60	8.59	57.0	-11.3	112.6
61	8.72	52.7	-10.1	93.3
62	8.84	47.8	-9.1	74.0
63	8.93	41.1	-8.0	51.3
64	9.07	33.3	-6.6	29.3
65	9.29	25.9	-4.4	12.7
66	**9.64**	**17.2**	**-0.9**	**-5.8**
67*	9.72	14.4	-0.1	-15.4
68*	9.70	13.1	-0.2	-21.9
69*	9.68	12.5	-0.4	-26.0
70*	9.62	12.0	-0.6	-31.2
71*	9.70	12.1	-0.2	-35.9
72*	9.63	106.3	0.4	462.9

Table 5.4 *Complete performance metrics per experiment, including rogue seasonality indices.*

(a) *Experimental clusters 1 & 2*

	$\overline{BW_o}$	$\overline{BW_I}$	$\Delta_O\%$	$\Delta_I\%$	$ITAE_o$	$ITAE_I$i	ϕ	ϕ_O	Φ_I
1	0.68	12.94	-2.25	25.14	12.96	131.86	2.43	1.94	2.93
2	0.68	12.41	-2.26	20.74	12.37	129.69	2.18	1.69	2.67
3	0.68	12.00	-2.27	17.42	11.83	128.41	2.00	1.50	2.49
4	0.68	11.27	-2.24	11.54	11.15	129.05	1.70	1.20	2.19
5	0.68	10.94	-2.20	8.93	10.86	134.23	1.57	1.08	2.06
6	0.68	10.40	-2.12	4.79	10.85	135.03	1.39	0.90	1.88
7	0.68	10.24	-2.08	3.56	10.79	43.65	1.34	0.85	1.82
8	0.68	9.79	-1.97	0.19	10.75	51.67	1.20	0.71	1.69
9	**0.68**	**9.50**	**-1.87**	**-2.00**	**10.72**	**162.06**	**1.12**	**0.63**	**1.60**
10*	0.68	9.13	-1.72	-4.72	10.68	178.49	1.02	0.54	1.50
11*	0.68	8.87	-1.60	-6.61	10.66	187.27	0.96	0.48	1.44
12*	0.68	8.59	-1.45	-8.66	10.62	204.40	0.90	0.42	1.37
13*	0.68	8.25	-1.24	-11.19	10.48	260.45	0.83	0.36	1.30
14*	0.69	7.80	-0.92	-14.48	10.46	264.69	0.75	0.30	1.20
15*	0.69	7.77	-0.89	-14.70	10.19	325.88	0.74	0.29	1.20
16*	0.69	7.20	-0.42	-18.80	9.57	410.55	0.67	0.25	1.10
17*	0.69	6.85	-0.08	-21.51	8.36	529.00	0.64	0.25	1.03
18*	0.69	6.71	0.06	-22.56	2.32	926.95	0.63	0.25	1.01
19	1.88	27.88	-4.84	90.35	28.37	160.29	6.75	6.28	7.21
20	1.89	25.31	-5.21	71.28	24.87	141.26	5.23	4.77	5.68
21	1.89	23.85	-5.48	61.27	23.00	132.42	4.21	3.77	4.64
22	1.88	21.59	-5.79	46.54	20.22	118.99	3.51	3.09	3.93
23	1.88	19.75	-5.88	35.14	18.77	112.35	3.00	2.60	3.41
24	1.88	18.47	-5.81	27.63	15.86	100.29	2.59	2.20	2.99
25	1.89	17.01	-5.58	19.45	14.79	96.23	2.23	1.85	2.61
26	1.89	15.66	-5.21	12.27	13.10	89.90	1.99	1.62	2.37
27	1.90	14.40	-4.69	5.85	12.02	86.39	1.78	1.41	2.14
28	1.92	13.21	-4.04	-0.01	10.78	83.50	1.51	1.15	1.86
29	**1.93**	**12.03**	**-3.23**	**-5.61**	**10.02**	**84.85**	**1.15**	**0.81**	**1.49**
30*	1.95	10.78	-2.17	-11.47	10.00	90.63	0.66	0.40	0.92
31*	1.98	9.69	-1.09	-16.70	10.00	104.77	0.59	0.36	0.81
32*	2.00	8.69	-0.03	-21.83	10.00	124.66	0.54	0.41	0.68
33*	2.01	8.15	0.50	-24.89	9.99	167.88	0.56	0.48	0.64
34*	2.02	7.62	0.90	-28.19	10.04	219.99	0.58	0.55	0.61
35*	2.02	7.20	1.11	-31.09	10.06	309.08	0.64	0.68	0.60
36*	2.02	10.63	1.15	2.17	9.37	479.68	0.73	0.84	0.61

(b) *Experimental clusters 3 & 4*

	$\overline{BW_o}$	$\overline{BW_I}$	$\Delta_O\%$	$\Delta_I\%$	$ITAE_o$	$ITAE_I$i	ϕ	ϕ_O	Φ_I
37	4.27	51.02	-6.24	170.05	55.90	248.30	4.65	4.17	5.14
38	4.21	45.23	-7.78	133.44	49.86	216.61	3.62	3.14	4.11
39	4.18	40.63	-8.71	106.17	43.53	185.83	3.14	2.66	3.62
40	4.17	36.98	-9.01	85.78	38.40	162.74	2.49	2.01	2.96
41	4.17	34.00	-8.93	70.04	33.83	143.27	2.02	1.55	2.48
42	4.19	31.25	-8.63	56.51	29.45	125.27	1.73	1.27	2.19
43	4.21	28.54	-8.12	44.07	26.06	112.15	1.44	0.99	1.89
44	4.23	26.65	-7.66	36.04	23.03	101.76	1.20	0.75	1.65
45	4.26	24.88	-7.15	28.87	20.48	92.80	1.00	0.56	1.44
46	4.29	22.58	-6.40	20.25	17.59	83.37	0.83	0.40	1.27
47	4.34	19.45	-5.11	9.40	13.76	72.48	0.69	0.27	1.11
48	**4.49**	**14.45**	**-2.21**	**-6.63**	**10.01**	**67.27**	**0.59**	**0.20**	**0.98**
49*	4.52	13.22	-1.37	-10.74	10.00	71.34	0.56	0.23	0.88
50*	4.58	11.36	-0.17	-17.95	10.00	82.78	0.56	0.31	0.81
51*	4.60	10.40	0.26	-22.61	10.00	104.40	0.57	0.37	0.77
52*	4.61	9.81	0.41	-26.20	10.24	148.80	0.59	0.44	0.73
53*	4.60	9.24	0.50	-30.34	11.09	222.40	0.61	0.51	0.71
54*	4.60	8.93	0.63	-33.12	12.30	467.90	0.63	0.40	0.87
55	8.98	101.06	-7.12	355.88	88.20	352.39	12.78	12.33	13.22
56	8.61	86.03	-11.36	267.52	81.80	311.19	9.51	9.08	9.94
57	8.52	75.48	-12.38	207.00	76.66	280.21	7.56	7.15	7.97
58	8.56	70.01	-11.53	176.65	69.40	245.97	6.67	6.28	7.07
59	8.54	62.78	-11.75	140.03	63.49	219.21	5.67	5.28	6.05
60	8.59	56.99	-11.25	112.57	56.13	189.79	4.93	4.55	5.32
61	8.72	52.68	-10.14	93.33	50.61	168.75	4.38	3.99	4.77
62	8.84	47.83	-9.09	73.98	44.80	149.37	3.74	3.33	4.15
63	8.93	41.11	-7.97	51.27	40.25	133.69	2.89	2.47	3.31
64	9.07	33.32	-6.61	29.34	29.91	102.92	2.02	1.59	2.45
65	9.29	25.88	-4.41	12.65	20.84	77.68	1.36	0.96	1.77
66	**9.64**	**17.24**	**-0.89**	**-5.81**	**11.71**	**58.85**	**0.71**	**0.49**	**0.92**
67*	9.72	14.35	-0.11	-15.42	10.00	62.83	0.55	0.47	0.64
68*	9.70	13.08	-0.24	-21.89	10.02	98.74	0.55	0.54	0.56
69*	9.68	12.49	-0.44	-25.99	10.40	121.89	0.58	0.61	0.55
70*	9.62	11.99	-0.58	-31.21	11.71	165.50	0.66	0.75	0.58
71*	9.70	12.05	-0.24	-35.89	16.77	353.71	0.91	1.09	0.74
72*	9.63	106.28	0.36	462.92	19.98	669.67	0.64	0.36	0.91

Table 5.5 *Performance metrics for seasonal demand streams. Experimental clusters 1 & 2.*

	Demand Stream 2					Demand Stream 3					Demand Stream 4				
	BW_o	BW_i	$\Delta_o\%$	$\Delta_I\%$	ϕ	BW_o	BW_i	$\Delta_o\%$	$\Delta_I\%$	ϕ	BW_o	BW_i	$\Delta_o\%$	$\Delta_I\%$	ϕ
1	0.56	8.75	-19.00	-15.32	1.64	2.42	62.56	249.09	505.08	8.72	1.17	24.70	69.22	138.91	4.24
2	0.57	8.39	-17.41	-18.37	1.47	2.27	57.47	226.61	459.24	7.49	1.17	24.36	68.89	137.07	3.90
3	0.58	8.11	-16.00	-20.67	1.35	2.15	53.72	209.43	425.49	6.63	1.17	24.13	68.57	136.06	3.63
4	0.60	7.60	-12.94	-24.74	1.15	1.93	47.25	178.83	367.75	5.26	1.17	23.79	67.94	135.48	3.15
5	0.62	7.37	-11.28	-26.54	1.06	1.84	44.47	165.24	343.01	4.71	1.16	23.67	67.59	135.79	2.95
6	0.64	7.01	-8.16	-29.41	0.94	1.69	40.19	143.78	304.97	3.91	1.16	23.53	66.82	137.08	2.63
7	0.64	6.89	-7.10	-30.26	0.91	1.65	38.93	137.44	293.90	3.69	1.16	23.51	66.62	137.82	2.54
8	0.67	6.59	-3.76	-32.57	0.82	1.53	35.63	120.28	264.56	3.12	1.15	23.44	65.50	139.84	2.30
9	**0.69**	**6.39**	**-1.22**	**-34.07**	**0.76**	**1.45**	**33.54**	**109.30**	**246.08**	**2.78**	**1.14**	**23.41**	**64.53**	**141.54**	**2.14**
10*	0.71	6.14	2.48	-35.92	0.70	1.36	31.02	95.85	223.70	2.38	1.13	23.39	62.94	144.07	1.96
11*	0.73	5.97	5.46	-37.18	0.66	1.29	29.32	86.73	208.61	2.13	1.12	23.38	61.56	146.09	1.83
12*	0.76	5.78	9.18	-38.54	0.62	1.23	27.53	76.97	192.62	1.87	1.11	23.33	59.48	148.02	1.70
13*	0.79	5.55	14.62	-40.17	0.59	1.15	25.38	65.27	173.41	1.57	1.08	23.23	56.21	150.23	1.53
14*	0.86	5.27	23.52	-42.20	0.56	1.05	22.74	50.72	149.49	1.22	1.04	22.90	50.13	151.21	1.31
15*	0.86	5.25	24.19	-42.33	0.56	1.04	22.57	49.79	147.95	1.20	1.04	22.87	49.65	151.17	1.30
16*	0.97	4.92	39.74	-44.55	0.55	0.92	19.50	33.09	119.78	0.84	0.96	21.87	38.74	146.45	1.02
17*	1.07	4.74	53.95	-45.68	0.54	0.85	17.64	23.01	102.27	0.66	0.90	20.70	29.61	137.31	0.84
18*	1.11	4.68	60.61	-46.00	0.54	0.83	16.95	19.33	95.71	0.61	0.87	20.12	25.84	132.30	0.77
19	1.41	18.74	-28.85	27.93	3.03	9.04	162.40	357.17	1008.82	17.55	1.85	27.17	-6.34	85.50	4.70
20	1.45	16.99	-27.35	15.01	2.41	7.40	124.91	270.86	745.38	11.44	1.85	25.16	-7.07	70.28	3.91
21	1.47	16.01	-26.34	8.24	2.12	6.70	109.55	235.46	640.68	9.23	1.85	24.07	-7.35	62.73	3.46
22	1.51	14.49	-24.23	-1.70	1.72	5.80	90.24	190.30	512.38	6.73	1.85	22.42	-7.61	52.18	2.78
23	1.56	13.24	-21.74	-9.36	1.43	5.17	77.19	158.77	428.24	5.21	1.85	21.12	-7.67	44.55	2.24
24	1.61	12.39	-19.43	-14.37	1.24	4.78	69.30	139.19	378.88	4.36	1.85	20.26	-7.59	39.98	1.90
25	1.68	11.42	-16.00	-19.77	1.04	4.37	61.23	118.61	329.97	3.54	1.85	19.32	-7.30	35.68	1.54
26	1.76	10.55	-11.81	-24.40	0.87	4.01	54.46	100.72	290.40	2.88	1.86	18.53	-6.85	32.79	1.25
27	1.87	9.74	-6.60	-28.39	0.73	3.68	48.55	84.41	256.85	2.32	1.87	17.87	-6.22	31.34	1.01
28	2.00	9.01	0.03	-31.79	0.62	3.37	43.17	68.85	226.83	1.84	1.89	17.37	-5.44	31.48	0.82
29	**2.18**	**8.34**	**9.07**	**-34.60**	**0.55**	**3.06**	**37.95**	**53.15**	**197.77**	**1.39**	**1.91**	**17.05**	**-4.44**	**33.80**	**0.67**
30*	2.46	7.72	23.30	-36.62	0.53	2.71	32.35	35.82	165.66	0.96	1.94	17.06	-3.11	40.10	0.56
31*	2.86	7.37	43.02	-36.66	0.53	2.40	27.29	19.96	134.55	0.63	1.96	17.54	-2.08	50.76	0.52
32*	3.45	7.41	72.37	-33.39	0.55	2.10	22.50	5.28	102.25	0.53	1.95	18.50	-2.25	66.35	0.52
33*	3.90	7.72	95.04	-28.90	0.58	1.95	19.84	-2.52	82.81	0.52	1.92	19.03	-3.95	75.34	0.52
34*	4.41	8.31	120.29	-21.75	0.64	1.80	17.25	-9.81	62.48	0.52	1.84	18.95	-7.89	78.55	0.53
35*	4.71	8.91	136.10	-14.67	0.73	1.69	15.18	-15.27	45.27	0.52	1.74	17.80	-12.92	70.43	0.53
36*	4.76	11.49	138.08	10.50	0.64	1.66	16.91	-16.98	62.55	0.57	1.70	19.59	-14.79	88.34	0.53

Table 5.6 *Performance metrics for seasonal demand streams. Experimental clusters 1 & 2.*

	Demand Stream 2					Demand Stream 3					Demand Stream 4				
	BW_o	BW_i	$\Delta_o\%$	$\Delta_I\%$	ϕ	BW_o	BW_i	$\Delta_o\%$	$\Delta_I\%$	ϕ	BW_o	BW_i	$\Delta_o\%$	$\Delta_I\%$	ϕ
37	3.09	34.22	-32.32	81.11	4.75	9.02	124.42	97.75	558.51	8.13	3.41	40.78	-25.23	115.84	6.76
38	3.10	30.33	-32.15	56.52	3.88	7.91	101.17	73.10	422.18	6.11	3.36	36.66	-26.39	89.21	5.31
39	3.14	27.24	-31.43	38.23	3.22	7.21	86.24	57.63	337.64	5.04	3.34	33.40	-27.09	69.49	4.24
40	3.20	24.80	-30.14	24.59	2.71	6.74	75.92	47.27	281.39	4.36	3.33	30.82	-27.34	54.84	3.48
41	3.28	22.82	-28.46	14.13	2.30	6.41	68.36	39.88	241.90	3.87	3.33	28.73	-27.34	43.67	2.95
42	3.37	21.01	-26.34	5.21	1.95	6.14	62.17	34.10	211.36	3.42	3.34	26.81	-27.15	34.26	2.55
43	3.51	19.24	-23.57	-2.86	1.62	5.92	56.65	29.05	185.95	2.95	3.36	24.94	-26.83	25.87	2.23
44	3.61	18.03	-21.17	-7.96	1.41	5.78	53.20	26.10	171.60	2.61	3.37	23.65	-26.49	20.71	2.03
45	3.74	16.91	-18.47	-12.39	1.23	5.66	50.15	23.38	159.78	2.27	3.39	22.45	-26.14	16.32	1.84
46	3.93	15.50	-14.23	-17.47	1.02	5.51	46.54	20.29	147.83	1.84	3.41	20.95	-25.56	11.54	1.59
47	4.29	13.67	-6.32	-23.11	0.77	5.31	41.94	15.98	135.93	1.29	3.45	18.98	-24.55	6.76	1.21
48	**5.40**	**11.43**	**17.77**	**-26.15**	**0.54**	**4.86**	**34.18**	**5.90**	**120.87**	**0.61**	**3.58**	**16.41**	**-22.03**	**6.07**	**0.65**
49*	5.89	11.23	28.46	-24.22	0.54	4.67	31.69	1.74	113.94	0.53	3.61	16.09	-21.18	8.60	0.57
50*	6.91	11.61	50.69	-16.13	0.59	4.25	26.69	-7.22	92.79	0.51	3.68	16.33	-19.80	17.98	0.51
51*	7.39	12.18	61.28	-9.39	0.73	3.97	23.25	-13.33	73.00	0.51	3.70	17.21	-19.37	28.04	0.51
52*	7.43	12.46	62.06	-6.26	0.89	3.77	20.69	-17.71	55.73	0.52	3.68	18.17	-19.70	36.77	0.51
53*	6.99	12.26	52.74	-7.62	1.02	3.56	17.83	-22.23	34.38	0.53	3.59	18.94	-21.43	42.76	0.52
54*	6.28	11.61	37.53	-12.98	1.02	3.40	15.54	-25.52	16.42	0.54	3.45	18.03	-24.35	35.11	0.53
55	6.26	67.77	-35.18	205.71	9.64	11.32	138.93	17.18	526.72	10.67	6.55	73.77	-32.26	232.80	12.38
56	6.09	57.69	-37.29	146.45	7.49	10.14	112.55	4.36	380.84	9.60	6.30	63.43	-35.19	170.97	9.16
57	6.12	50.60	-37.06	105.80	5.91	9.45	95.07	-2.77	286.65	8.93	6.23	56.15	-35.92	128.36	7.11
58	6.22	46.93	-35.70	85.46	5.02	9.17	86.29	-5.27	240.98	8.71	6.26	52.36	-35.35	106.89	6.29
59	6.32	42.12	-34.65	61.05	3.99	8.83	76.40	-8.70	192.08	8.19	6.24	47.37	-35.53	81.09	5.62
60	6.49	38.30	-32.95	42.88	3.23	8.63	68.92	-10.76	157.10	7.49	6.27	43.36	-35.22	61.76	5.40
61	6.71	35.50	-30.78	30.28	2.70	8.55	63.49	-11.83	132.98	6.75	6.35	40.39	-34.50	48.22	5.44
62	6.99	32.39	-28.16	17.80	2.19	8.48	57.97	-12.79	110.86	5.57	6.44	37.06	-33.82	34.79	5.37
63	7.35	28.14	-24.24	3.54	1.66	8.38	51.28	-13.62	88.69	3.81	6.49	32.48	-33.06	19.51	4.76
64	7.92	23.37	-18.45	-9.27	1.19	8.33	44.35	-14.24	72.17	2.18	6.59	27.23	-32.14	5.71	3.52
65	8.77	19.16	-9.75	-16.56	0.84	8.35	38.52	-14.07	67.72	1.13	6.74	22.35	-30.59	-2.68	2.19
66.00	**10.77**	**15.98**	**10.74**	**-12.74**	**0.59**	**8.25**	**32.07**	**-15.23**	**75.13**	**0.53**	**7.01**	**17.38**	**-27.89**	**-5.09**	**0.77**
67*	11.57	15.97	18.90	-5.87	0.75	7.89	28.31	-18.85	66.82	0.50	7.10	16.67	-27.01	-1.78	0.54
68*	11.34	15.70	16.64	-6.23	0.93	7.56	25.33	-22.25	51.31	0.51	7.11	17.21	-26.87	2.80	0.51
69*	10.92	15.23	12.22	-9.74	0.92	7.35	23.36	-24.42	38.40	0.52	7.10	17.99	-26.95	6.59	0.51
70*	10.16	14.41	5.00	-17.28	0.81	7.08	20.87	-26.89	19.77	0.53	7.04	19.34	-27.30	11.01	0.52
71*	9.32	13.56	-4.17	-27.83	0.68	6.90	18.39	-29.05	-2.17	0.59	6.97	20.33	-28.37	8.16	0.54
72*	8.83	76.14	-8.04	303.29	0.51	6.74	80.13	-29.73	324.39	1.45	6.81	82.92	-29.09	339.18	0.57

Appendix B
The Influence of Structural Parameters

In this appendix, we present frequency plots and bullwhip plots computed with different structural parameters. Each figure shows the plots corresponding to a fixed Cover (C) and a varying exponential Smoothing Coefficient (α) and Lead Time (L). Figures B.2–B.4 show the frequency plots of orders (with $C = 1, 3, 6$, respectively); Figures B.5–B.7, the frequency plots for inventories; Figures B.8–B.10, the bullwhip plots for orders; and Figures B.11–B.12, the bullwhip plots for inventories.

Each figure contains 9 sub-figures arranged in a $3x3$ matrix. Each row contains plots with fixed α and $L = 1, 5, 10$; and each column, plots with fixed L and $\alpha = 0.1, 0.3, 0.5$. The experimental design for the frequency plots is shown in Figure B.1.

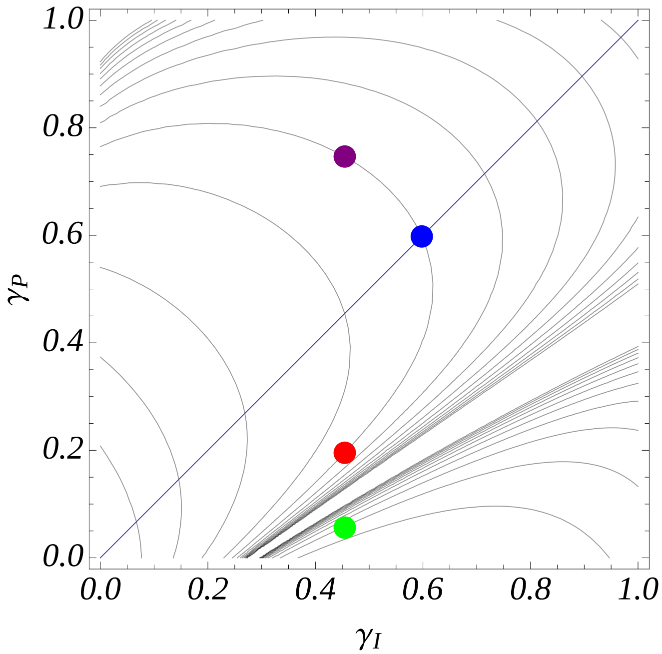


Figure B.1 *Experimental design. The color dots show 4 different parametrizations of the behavioral parameters. The parametrizations correspond to the same stationary performance* ($\mathrm{BW}_O = 6$)*. The blue line highligts the DE-diagonal.*

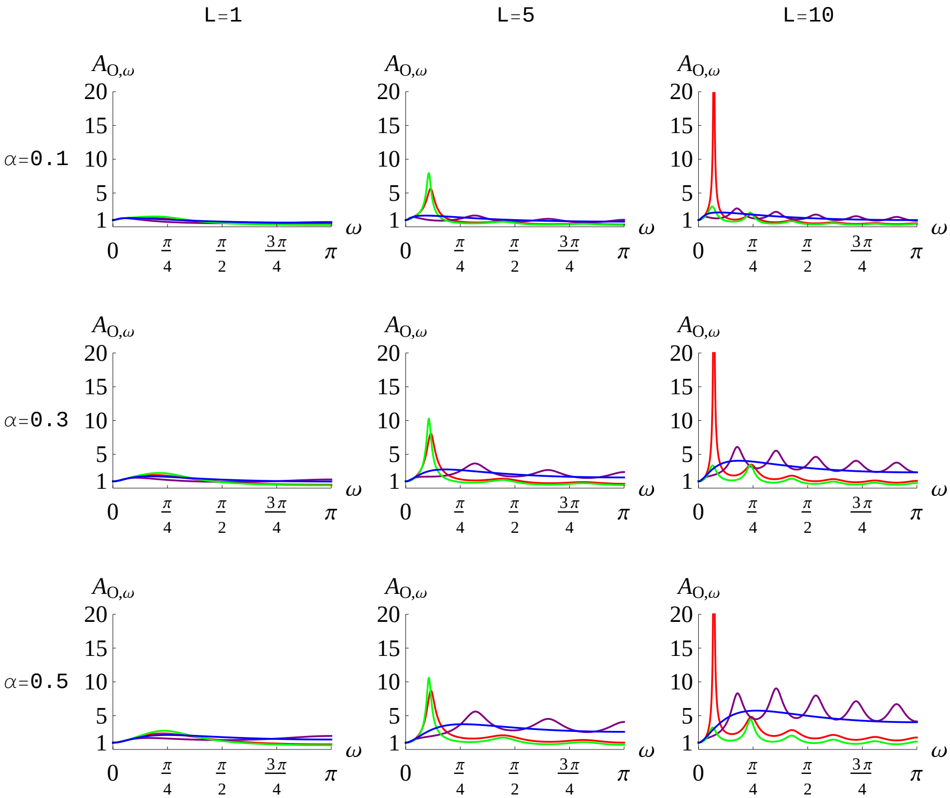


Figure B.2 *Frequency plots of orders corresponding to the experimental design of Figure B.1 and coverage $C = 1$. We see that the response of systems along the DE-diagonal (blue) maintain its flatness across variations in structural parameters; systems below the DE-diagonal (red) exhibit a low frequency peak; and systems above the DE-diagonal (purple) exhibit multiple frequency peaks, which depend on the lead time. Increasing the exponential smoothing parameter α increases the amplification at all frequencies. Lead time L changes affect the response of the system, the amount, position, and magnitude of frequency peaks appears to be associated with lead time, except for systems along the DE-line (blue).*

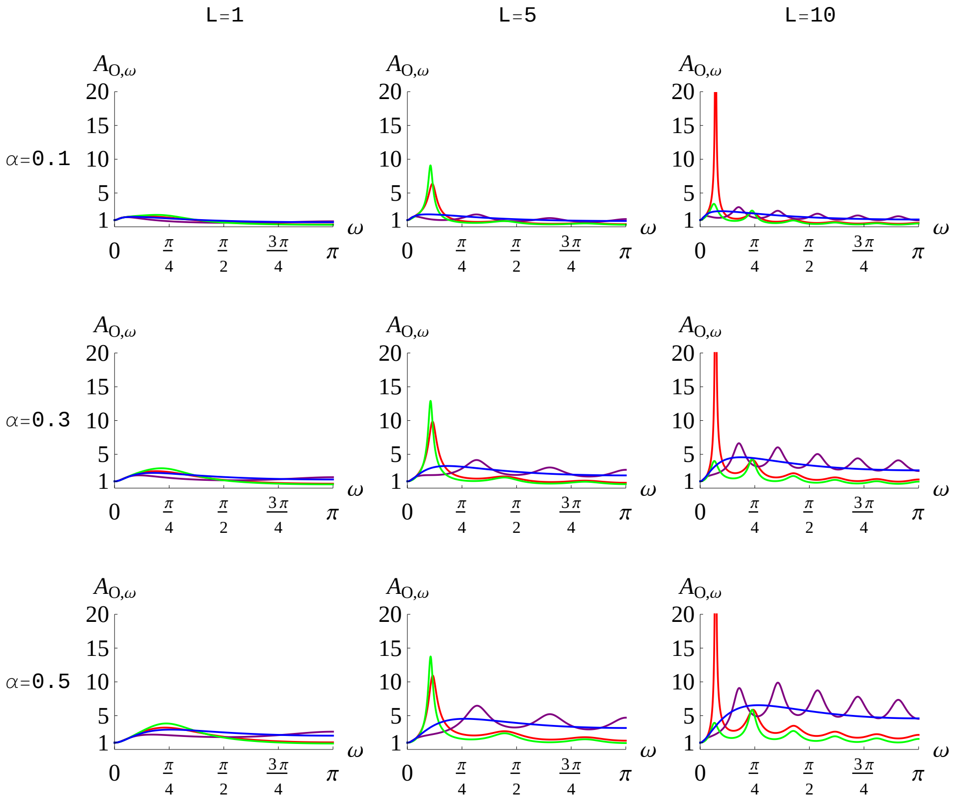


Figure B.3 *Frequency plots of orders corresponding to the experimental design of Figure B.1 and coverage $C = 3$. We see that the response of systems along the DE-diagonal (blue) maintain its flatness across variations in structural parameters; systems below the DE-diagonal (red) exhibit a low frequency peak; and systems above the DE-diagonal (purple) exhibit multiple frequency peaks, which depend on the lead time. Increasing the exponential smoothing parameter α increases the amplification at all frequencies. Lead time L changes affect the response of the system, the amount, position, and magnitude of frequency peaks appears to be associated with lead time, except for systems along the DE-line (blue).*

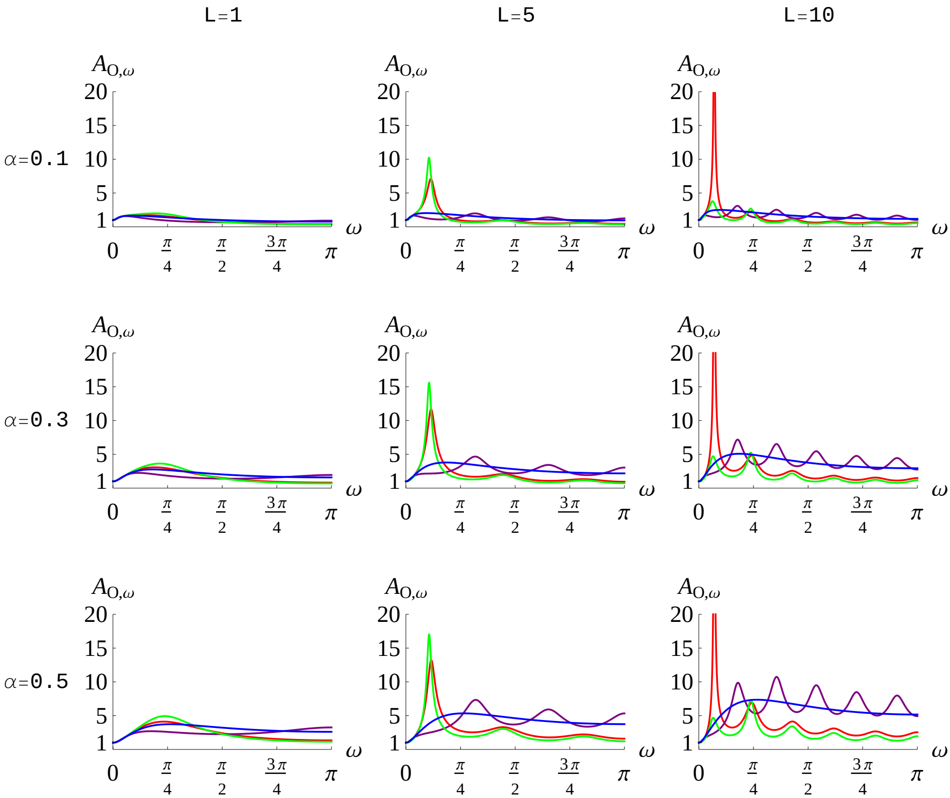


Figure B.4 *Frequency plots of orders corresponding to the experimental design of Figure B.1 and coverage $C = 6$. We see that the response of systems along the DE-diagonal (blue) maintain its flatness across variations in structural parameters; systems below the DE-diagonal (red) exhibit a low frequency peak; and systems above the DE-diagonal (purple) exhibit multiple frequency peaks, which depend on the lead time. Increasing the exponential smoothing parameter α increases the amplification at all frequencies. Lead time L changes affect the response of the system, the amount, position, and magnitude of frequency peaks appears to be associated with lead time, except for systems along the DE-line (blue).*

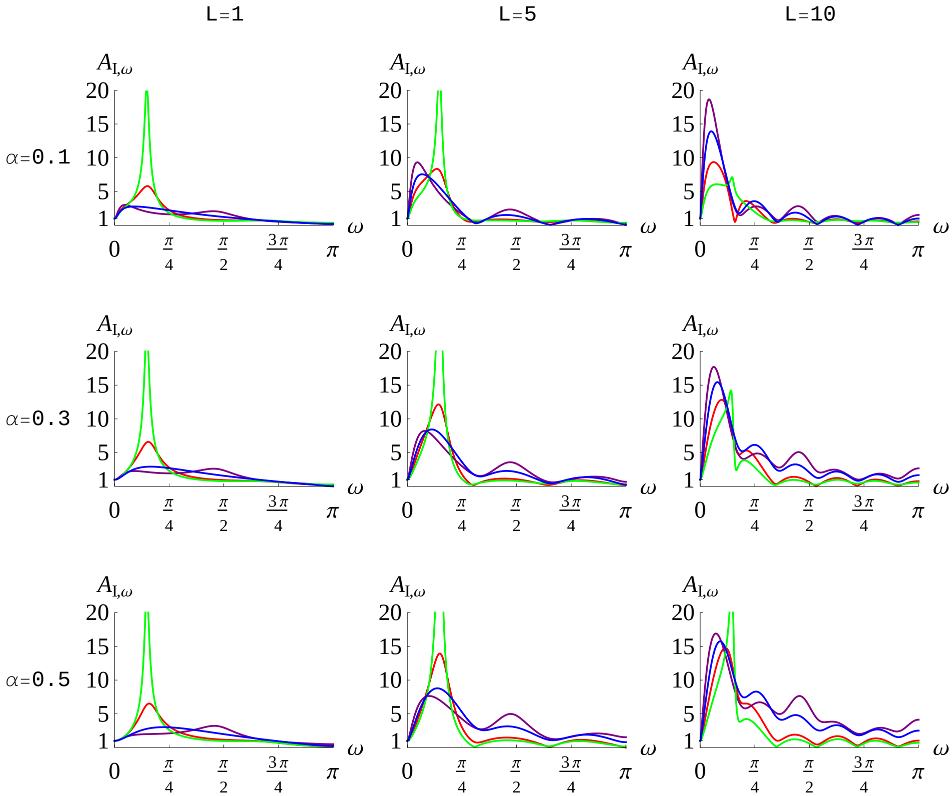


Figure B.5 *Frequency plots of inventories corresponding to the experimental design of Figure B.1 and coverage $C = 1$. We see that the lead time L changes affect the response of the system. The the amount, position, and magnitude of frequency peaks appears to be associated with lead time for all systems (including the DE-line). Systems below the DE-diagonal (red), and systems above the DE-diagonal (purple) exhibit multiple frequency peaks, which depend on the lead time. Increasing the exponential smoothing parameter α increases the amplification at all frequencies.*

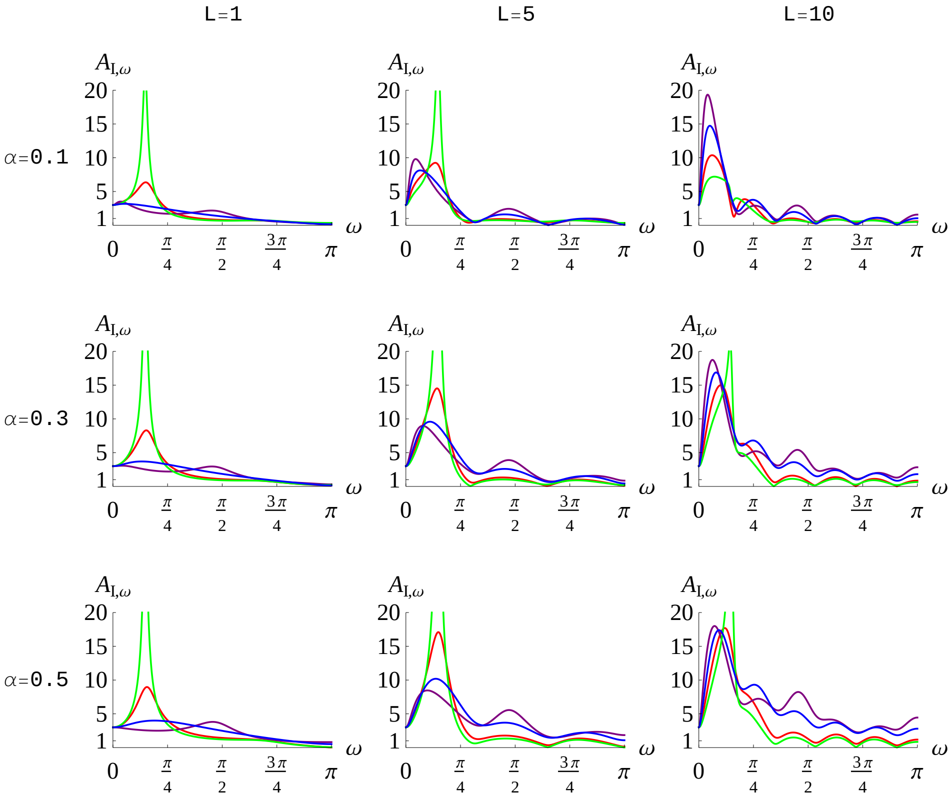


Figure B.6 *Frequency plots of inventories corresponding to the experimental design of Figure B.1 and coverage $C = 3$. We see that the lead time L changes affect the response of the system. The the amount, position, and magnitude of frequency peaks appears to be associated with lead time for all systems (including the DE-line). Systems below the DE-diagonal (red), and systems above the DE-diagonal (purple) exhibit multiple frequency peaks, which depend on the lead time. Increasing the exponential smoothing parameter α increases the amplification at all frequencies.*

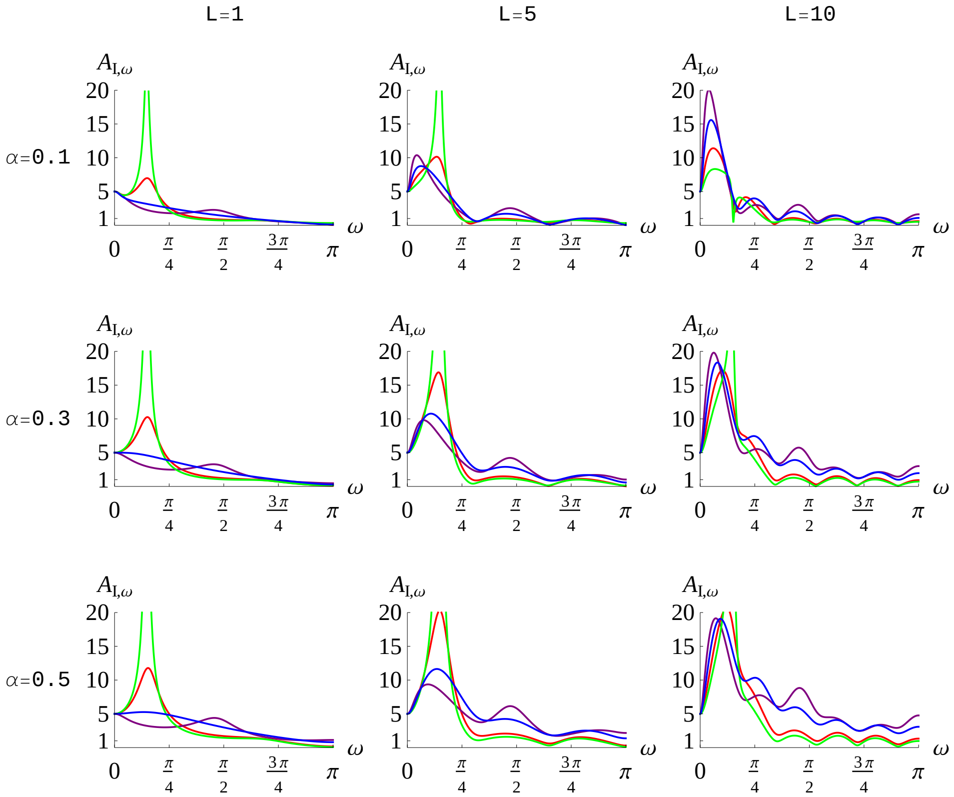


Figure B.7 *Frequency plots of inventories corresponding to the experimental design of Figure B.1 and coverage $C = 6$. We see that the lead time L changes affect the response of the system. The the amount, position, and magnitude of frequency peaks appears to be associated with lead time for all systems (including the DE-line). Systems below the DE-diagonal (red), and systems above the DE-diagonal (purple) exhibit multiple frequency peaks, which depend on the lead time. Increasing the exponential smoothing parameter α increases the amplification at all frequencies.*

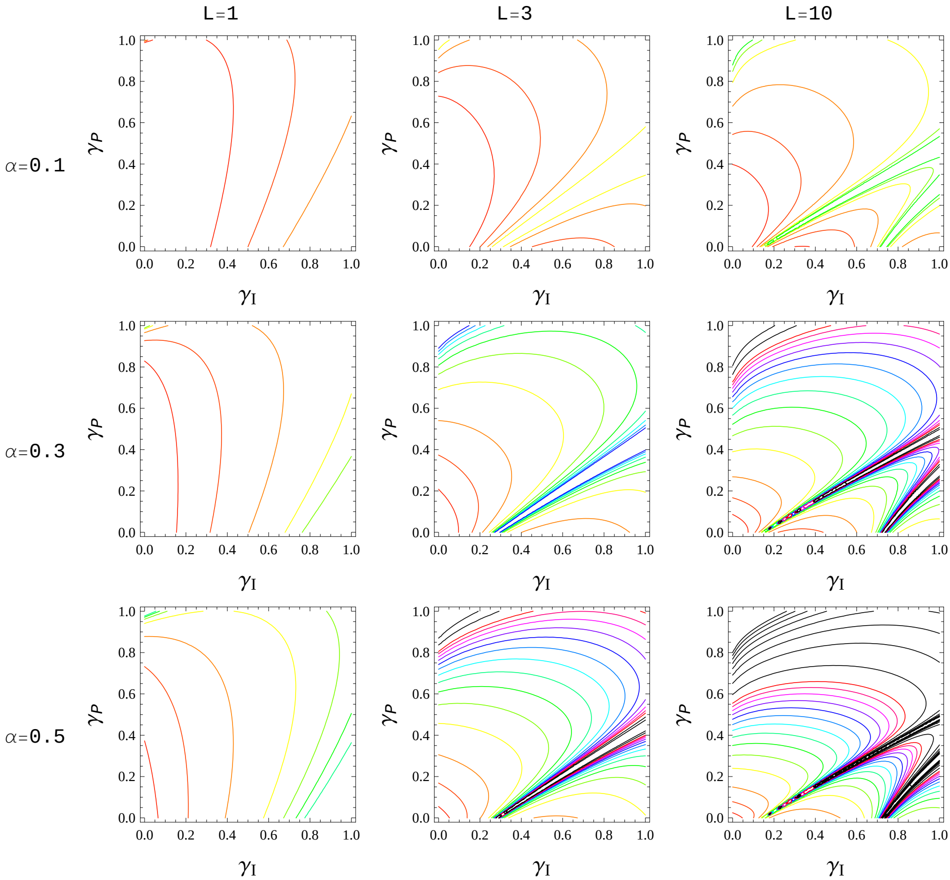


Figure B.8 *Bullwhip plots for orders for* C = 1. *The hue of each contour represents the value of the bullwhip, starting with dark orange* ($\mathrm{BW}_O = 0.5$) *and increasing towards dark red* ($\mathrm{BW}_O = 24$)*. Black contours indicate* $\mathrm{BW}_I \geq 30$ *and are illustrated in increments of 10.*

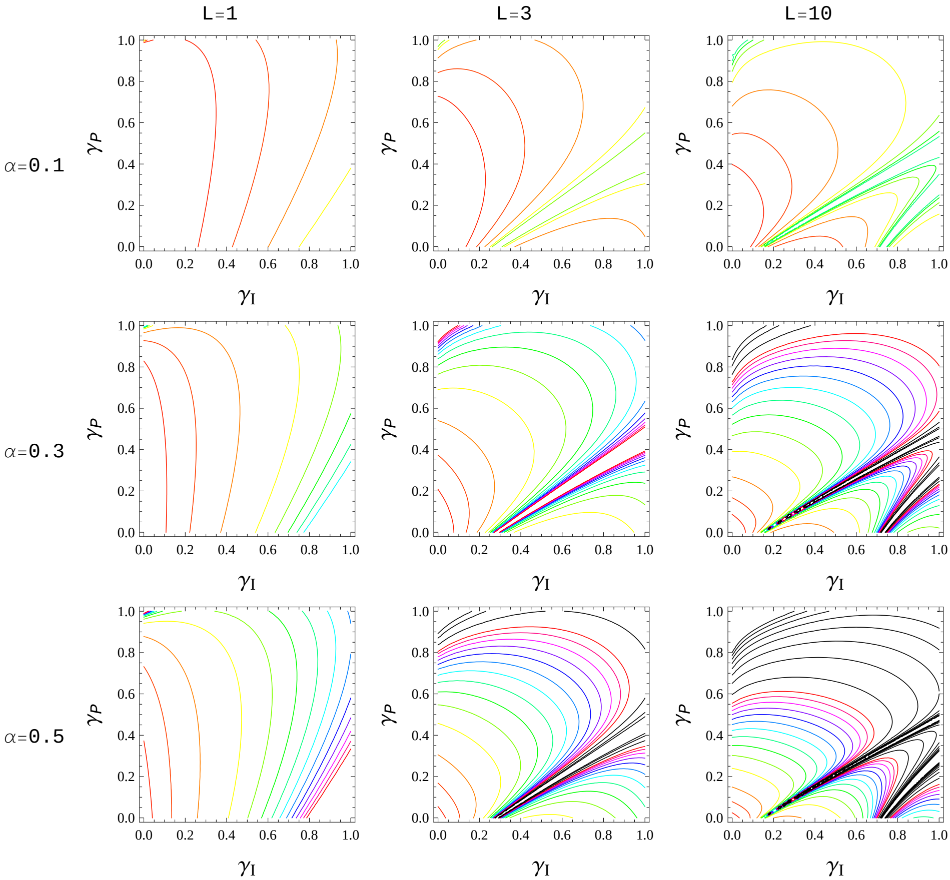


Figure B.9 *Bullwhip plots for orders for $C = 3$. The hue of each contour represents the value of the bullwhip, starting with dark orange ($\mathrm{BW}_O = 0.5$) and increasing towards dark red ($\mathrm{BW}_O = 24$). Black contours indicate $\mathrm{BW}_I \geq 30$ and are illustrated in increments of 10.*

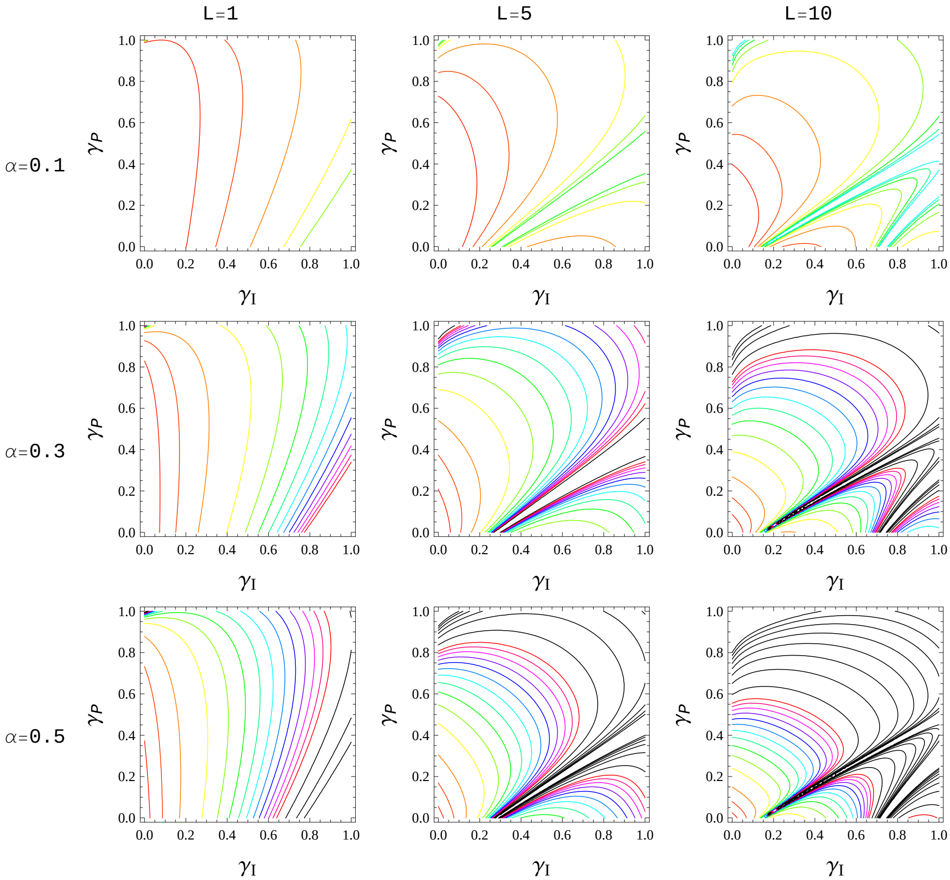


Figure B.10 *Bullwhip plots for orders for $C = 6$. The hue of each contour represents the value of the bullwhip, starting with dark orange ($\mathrm{BW}_O = 0.5$) and increasing towards dark red ($\mathrm{BW}_O = 24$). Black contours indicate $\mathrm{BW}_I \geq 30$ and are illustrated in increments of 10.*

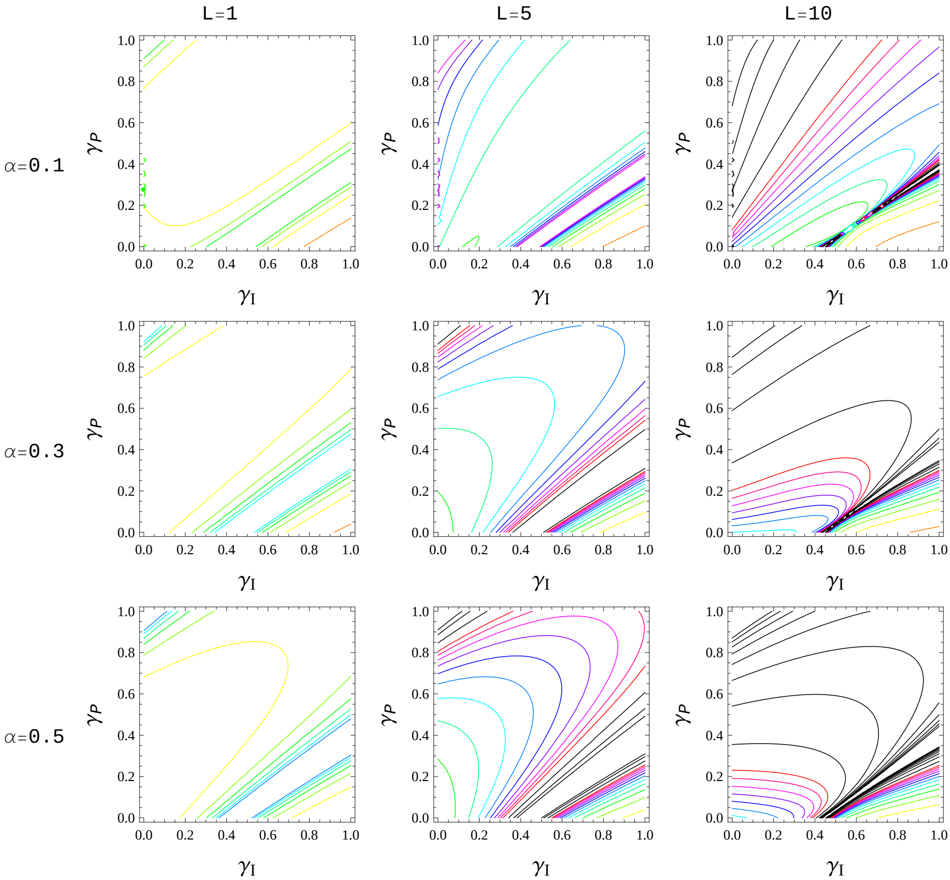


Figure B.11 *Bullwhip plots for Inventories for $C = 1$. The hue of each contour represents the value of the bullwhip, starting with dark orange ($\mathrm{BW}_I = 0.5$) and increasing towards dark red ($\mathrm{BW}_I = 24$). Black contours indicate $\mathrm{BW}_I \geq 30$ and are illustrated in increments of 10.*

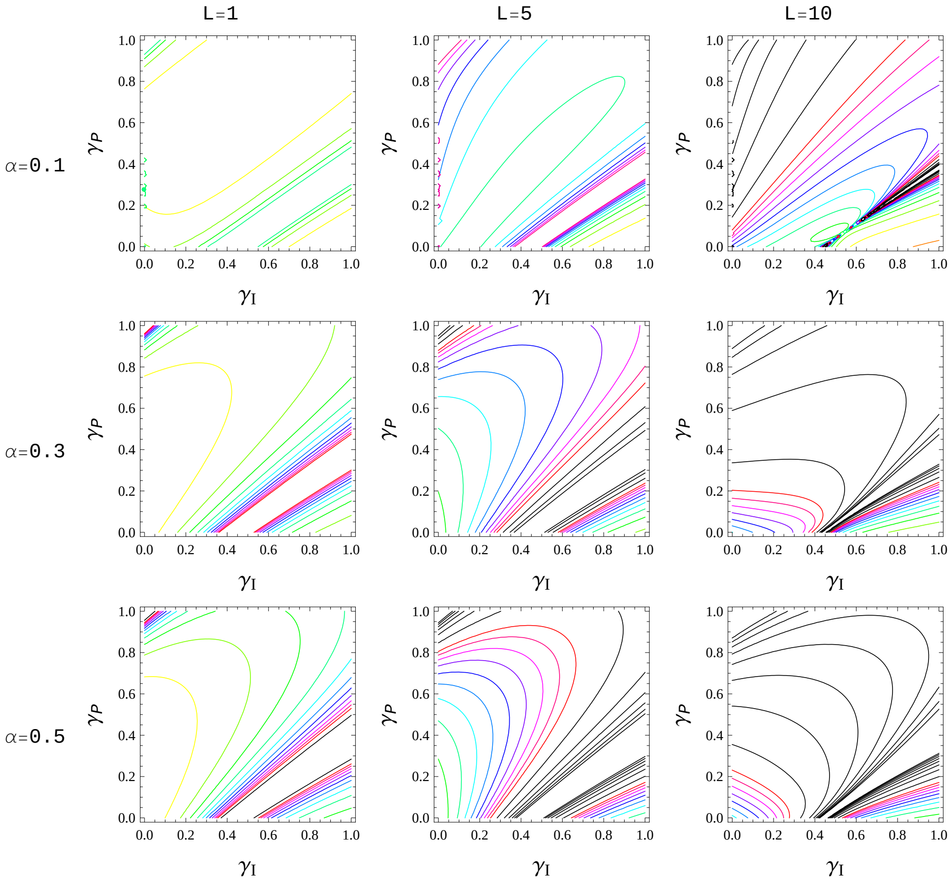


Figure B.12 *Bullwhip plots for Inventories for* $C = 3$. *The hue of each contour represents the value of the bullwhip, starting with dark orange* ($\mathrm{BW}_I = 0.5$) *and increasing towards dark red* ($\mathrm{BW}_I = 24$). *Black contours indicate* $\mathrm{BW}_I \geq 30$ *and are illustrated in increments of 10.*

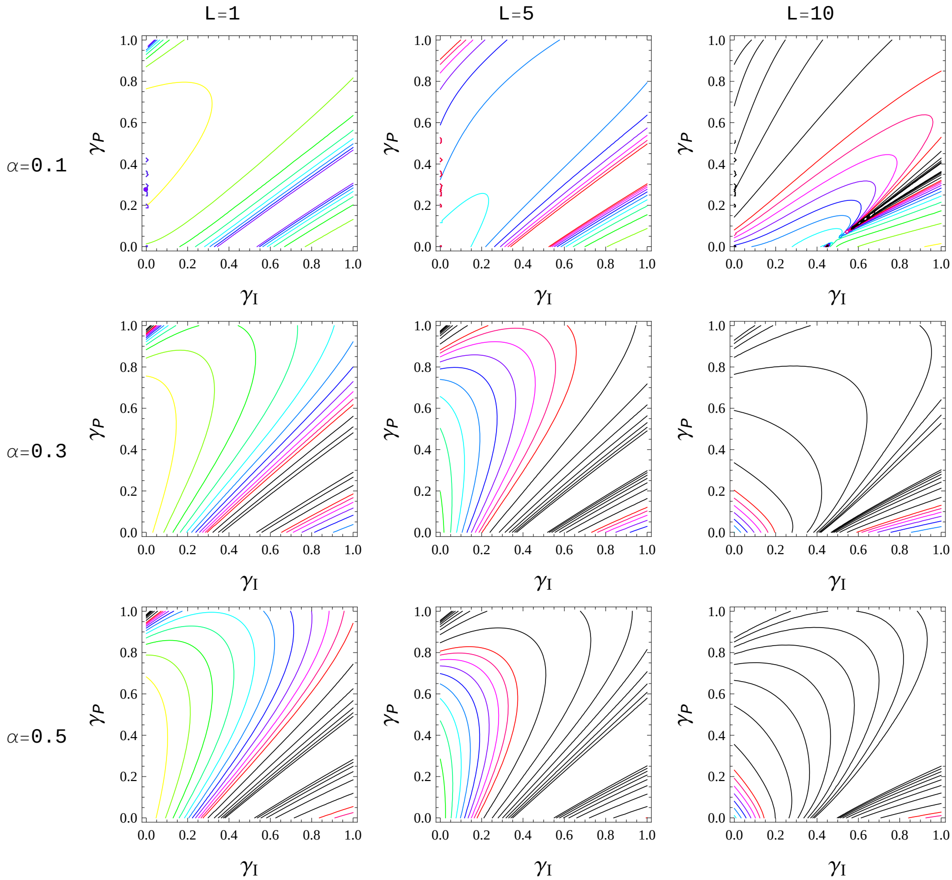


Figure B.13 *Bullwhip plots for Inventories for* $C = 6$*. The hue of each contour represents the value of the bullwhip, starting with dark orange* ($\mathrm{BW}_I = 0.5$) *and increasing towards dark red* ($\mathrm{BW}_I = 24$)*. Black contours indicate* $\mathrm{BW}_I \geq 30$ *and are illustrated in increments of 10.*

Although I cannot lay an egg,
I am a very good judge of
omelettes.

George Bernard Shaw

Chapter 6

Inventories and the Bullwhip: a Chicken and Egg Situation

Inventories, as we saw in Chapter 3, can play a large role in business cycles. However, as we saw in Chapter 2, the causal relationship is not entirely understood; do inventory cycles generate business cycles, or vice versa? In operations management theory, modeling choice pre-determines the role of inventories. Production smoothing models dictate that inventories follow production decisions; Order-Up-To models dictate that inventories drive production decisions. Empirical observations, however, show that inventories are both a cause and an effect of production decisions.

In this chapter, we hypothesize that –in real life– target inventory levels are dynamic and are adjusted in response to not only changes in demand forecast but also fluctuating cost considerations. We develop a two-echelon structural model of inventory decisions using financial data for 6040 unique supplier-customer dyads for the years 1984–2013 to investigate downstream inventory adjustments and their influence on upstream firms. The model shows that suppliers react to arbitrary downstream inventory adjustments over and above the demand changes, revealing a new explanation for transient shocks getting amplified upstream. Our results show that inventory cost ratios are dynamic, and support the hypothesis that they follow economic and financial sentiment such as liquidity considerations and GDP growth rates.

6.1 Introduction

The importance of inventories in the economy has long been understood. Economists agree that fluctuations in inventory investment account for more than a third of the quarterly change in U.S. GDP (Fitzgerald, 1997)[1]. Blinder (1990) famously remarked: "Business cycles are, to a surprisingly large degree, inventory cycles". Yet, the precise role of inventories in these cycles is still an open question. Do inventory cycles cause business cycles or is it the other way around?

At a smaller scale, the role of inventories in the day-to-day operations of a firm is, to a large extent, determined by its production strategy[2]. When convex production costs dominate, firms benefit from maintaining production as constant as possible. As a result, such firms use inventories as a buffer to absorb demand uncertainty. This is the intuition behind production smoothing.

When inventory-related costs dominate, on the other hand, firms benefit from maintaining a stable inventory level. In these firms, inventories drive production decisions with the objective of minimizing inventory variability around a pre-defined target level. This is the intuition behind Order-Up-To policies.

In this view, inventories can be classified as either an adjustment variable (i.e. the consequence of production decisions) or as a decision variable (i.e. the driver of production decisions). The proverbial "chicken or egg" situation. Reality, however, is more subtle. Neither pure production smoothing nor Order-Up-To models fully explain empirical observations. As a consequence, hybrid models have been developed: Smoothing models that incorporate inventory-driven adjustments (Blinder, 1990; Ramey and West, 1999) as well as generalized Order-Up-To models that incorporate a smoothing component (Dejonckheere et al., 2003; Chen and Lee, 2009). These models, however, assume that relative costs are constant in time; modeled firms do not change their smoothing strategy nor their inventory targets[3].

In this chapter, we argue that target inventory levels, in practice, are dynamic. Firms systematically adjust their targets responding (among other causes) to

[1]Depending on time period analyzed, estimations of how much inventory fluctuations account for post-war GDP fluctuations go from 44% to 87%, (Blinder and Maccini, 1991; Wang and Wen, 2009).

[2]Even though we use the term "production", the discussion also applies to non-manufacturing firms where "production" is replaced by "orders".

[3]Several models do however, compute target inventory levels as a function of a demand forecast that is updated with each subsequent demand observation. In these models, changes in the demand forecast result in changes in the target inventory level. However, it is not a systematic one; the underlying target inventory, expressed as a function of the demand forecast, remains constant.

fluctuating cost considerations. As a consequence, the customer demand faced by upstream firms depends directly on downstream inventory decisions.

This view is motivated by a series of recent developments. In Chapter 3, we discussed one way in which inventory decisions are able to drive the dynamics of entire supply chains. We conducted a firm-level study using the 2008 financial crisis as a natural experiment. We found our observations to be consistent with the hypothesis that inventories are used as an instrument of liquidity in times of need. Drastic reductions in inventory investment free up much needed cash. Similarly, Pesch and Hoberg (2013) used secondary data to show that inventories are used as a significant source of liquidity by firms in financial distress, independent of global crises.

Burns and Sivazlian (1978) analyzed inventory adjustments from a different perspective, using numerical experimentation in a serial supply chain. Without assumptions regarding the underlying motives, they show that inventory adjustments executed by downstream echelons trigger transient changes in their orders that cause an overreaction in upstream echelons because suppliers cannot distinguish transient from permanent changes in demand. They identify this "false ordering" phenomenon as a purely structural issue, product of the delays inherent to the transmission of information, and propose a modified ordering rule that takes this into account as a way of solving it. Intuitively, their solution implies separating the component of incoming orders that corresponds to downstream inventory fluctuations from the component that corresponds to "actual" demand, and reacting only to the latter.

More recently, the interaction between different echelons in a supply chain has been the focus of the information-sharing literature. The objective of this stream of the literature is, as Cachon and Fisher (2000) put it, to test "the general belief within industry that capturing and sharing real-time demand information is the key to improved supply chain performance". Insights from the analytical information-sharing literature indicate that, in general, upstream firms benefit the most from the sharing of information among supply chain players, and that the benefits depend on the type of demand observed downstream (and thus, the amount of non-inferable information that can be obtained through sharing). Most studies in this stream, however, assume rational firms with actions modeled through Order-Up-To policies, the knowledge of which is also shared among a supply chain. Chen and Lee (2009) drop the latter assumption, noting that in practice, it is rather rather bold to assume that firms share details about their ordering policies. They advocate for the sharing of order projections as it eliminates guesswork from the part of the suppliers. In fact, they note that this allows suppliers to separate order uncertainty from order variability.

Separating order uncertainty from order variability, according to Bray and Mendelson (2013), is the key to disentangle the effects of production smoothing and the Bullwhip Effect. They argue that the goal of production smoothing is to protect against order uncertainty, not against order variability—and thus, that empirical measures of production smoothing should not benchmark order variability against demand variability but against what the theoretical order variability would be in the absence of production adjustment costs.

In this chapter, we use secondary data at two levels of a supply chain (supplier/customer pairs) to investigate the causes of downstream inventory decisions, and their influence on upstream orders. Using firm level data, we find evidence of an overreaction of upstream firms to downstream inventory adjustments, and track down systematic factors that drive fluctuations of safety inventories. Methodologically, we use an econometric model specification to estimate the influence of inventory changes in orders, and structural modeling of the cost factors to estimate the systematic adjustment of inventory buffers.

The contribution of this chapter to the literature is threefold: First, we identify the portion of order uncertainty that corresponds to downstream safety stock changes and analyze its influence on upstream order generation. Second, we extend the structural modeling methodology of Olivares et al. (2008) to estimate the cost ratios used by firms in a multi-period setting. Finally, we show evidence of a systematic adjustment of target inventories that follows economic and financial conditions.

The remainder of the chapter is organized as follows. In Section 6.2, we introduce the data used and detail the construction of supplier/customer firm pairs. We present the econometric inventory model and related hypothesis in Section 6.3, the structural cost model and related hypothesis in Section 6.4. We follow in Section 6.5 with the results of our analysis, and conclude in Section 6.6.

6.2 Data

To quantify the effect of downstream inventory decisions on upstream firms, the first step we take is to collect firm-level data and define explicit relationships between upstream and downstream firms. To do so, we adopt an approach pioneered by the financial community. In this approach we use the disclosure of major customers, contained in the annual financial statements of public firms, to identify explicit relationships between firms. Recent examples of such an approach include Fee and Thomas (2004), who use customer-supplier relationship data to quantify the effect of horizontal mergers in post-merger operating performance; Fee et al. (2006), who investigate the customer-supplier relationships to determine the conditions under which customers own equity of their suppliers; and Cohen and Frazzini (2008), who study whether stock returns can be predicted through knowledge of economic links between companies. In §6.2.1 we detail the methodology we use to construct supplier-customer pairs, and in §6.2.2 we summarize the collection of the remaining financial data.

6.2.1 Customer-Supplier Firm Pairs

The regulation of the Statement of Financial Accounting Standards (SFAS) No. 131 requires firms to disclose any customer that represents at least 10% of the revenues for a given fiscal year. This information is included in Compustat's *customer segment database*, which we access through Wharton Research Data Services (WRDS) of the University of Pennsylvania. We extract the identity of supplier-customer firm pairs for the $1976 - 2012$ period.[4] Firms, however, report the identity of their customers as a plain-text string in a non-standardized way. Therefore, spelling mistakes exist in the data, as well as spelling variations and abbreviations that vary across suppliers and periods. We apply a three-step procedure, inspired by the financial research literature (Fee and Thomas, 2004), to map the reported plain-text customer names to Compustat's unique customer identifier keys. In the first step, we perform a 1-1 match between the reported company names and the Compustat company name and assign the corresponding customer identifier key to successful matches.

In the second step, we apply a partial string matching algorithm, based on the normalized Levenshtein distance, to the remaining data[5].

The partial string matching algorithm calculates the normalized Levenshtein distance between all potential combinations of reported and Compustat company names, and flags them as a potential match when $\overline{lev}_{a,b} < 0.25$. Potential matches are then manually checked against (a) possible ambiguities (in cases where any ambiguity exists –such as for example a match for 'continental', a name that appears in numerous companies– the potential matches are dropped), and (b) business segments, by confirming that the business segment reported by the supplier is compatible with the business segment of the matched customer.

[4] Regulation SFAS 131, issued by the Financial Accounting Standards Board in 1997, has been effective for fiscal years beginning after December 15th 1997. The customer disclosure requirements in this regulation carried over from regulation SFAS 14, effective since December 1976.

[5] The Levenshtein distance between strings a and b ($lev_{a,b}$) is defined as *The smallest number of insertions, deletions, and substitutions required to change string a into string b* (Levenshtein, 1966). The normalized Levenshtein distance ($\overline{lev}_{a,b}$) is defined as $lev_{a,b} / \min(|a|, |b|)$, where $|i|$ is the length of string i.

Finally, we sort all remaining unmatched observations by the frequency of appearances of the reported customer name and manually match the top 16 firms (these represent $\sim$ 16% of the data).

Compustat's *customer segment database* contains a discontinuity in the year 2006: No data is available for firms whose fiscal year ends in Q4 2006, causing the number of records for the year to be abnormally low.[6] To overcome this, we assume that companies that are linked both in 2005 and 2007 are also linked in 2006 and update the links accordingly.

As a reference, Figure 6.1 shows the number of records per quarter and the quarter-on-quarter change in US GDP for the entire period. Because the 1976-1984 period is marked by a large instability in the GDP time series and relatively few data-points, we drop the observations from 1976 until 1983 from the dataset. The final firm pair database includes 24825 unique yearly pairs.

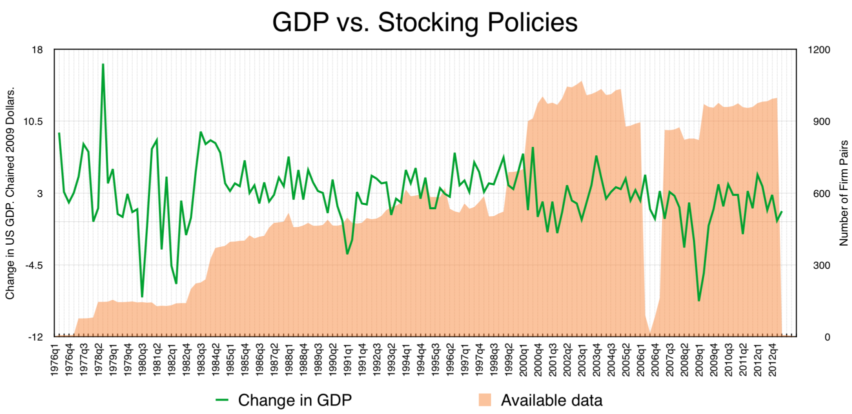


Figure 6.1 *Available data per quarter and GDP variation*

6.2.2 *Customer and Supplier Financial Data*

To take advantage of the reporting frequency of the *quarterly fundamentals database*, we assume that each of the firm pairs we have defined are involved in trading during the four quarters of the reported fiscal year, and thus populate the firm pair database with financial information from said database. Following Cachon et al. (2007), we use cost of goods sold (COGS) as a proxy for demand (d), production (q) as a proxy for orders, and calculate production with the balance equation $q_t = d_t + i_t - i_{t-1}$, where i_t is the reported total inventory at the end

[6]This is due to significant changes made to Compustat's databases in the year 2006.

of the period t. We delete observations with negative, or missing, values for inventories and/or sales from the sample.

Since fiscal year endings vary across firms, we pair customers and suppliers independently using calendar dates. The final sample contains 77886 quarterly observations, representing 6040 unique supplier-customer relationships between 1984 and 2013. Table 6.1 shows summary statistics for our dataset.

The firm size distribution is calculated on the basis of average COGS over the entire sample. We can see that the nature of our supplier-customer pairing strategy, where we can only identify customers that represent more than 10% of a firm's revenues, biases the sample towards large customers with a large number of relatively smaller suppliers. We analyze the influence of this bias on our results in Appendix C, presented at the end of this chapter.

Table 6.1 *Summary Statistics*

	Min	Max	Mean	SD	1st Quartile	Median	3rd Quartile
Number of suppliers in the sample per quarter	5	711	501	142	391	533	609
Number of customers in the sample per quarter	4	349	220	110	101	277	321
Frequency-Weighed link duration (quarters)	1	140	14	15	4	8	17
Supplier size percentile	0.01	0.99	0.70	0.82	0 .24	0.42	0.62
Customer size percentile	0.01	0.99	0.99	0.99	0 .91	0.98	0.99
Number of customers per supplier	1	27	3.78	3.20	2	3	5
Number of suppliers per customer	1	351	102.1	112.1	11	53	203

Table 6.2 shows additional sales and inventory statistics of the firms contained in our dataset, grouped according to their 2-digit NAICS code. Note that while we constrain the suppliers to the manufacturing sector, customers belong to a range of industries. Nevertheless, manufacturing, retail, and wholesale customers make up the vast majority of the sample: 58%, 19% and 10% respectively.

6.3 *Replenishment Model*

In this section, we use a non-stationary-demand inventory model to motivate hypotheses related to the influence of downstream inventory changes on upstream orders. In §6.3.1 we present the single echelon inventory model. We derive the hypotheses in §6.3.2, detail the 2-echelon model in §6.3.3, present a forecasting model in 6.3.4, and finally, present the econometric specification of the model in §6.3.5.

Table 6.2 *Summary statistics by NAICS code for suppliers (top) and customers (bottom)*

NAICS Code	Frequency	COGS				Total Inventory			
Supplier		Mean	25th Percentile	Median	75th Percentile	Mean	25th Percentile	Median	75th Percentile
31	8070	294.7	17.8	56.5	217.4	237.1	15.3	53.8	188.6
32	17120	181.4	5.9	20.8	109.6	224.9	1.6	15.1	100.7
33	52696	252.6	7.3	30.7	135.1	156.9	7.3	27.2	98.3
Customer									
11	18	662.0	350.0	555.5	710.0	1387.0	1130.0	1298.5	1387.0
21	323	2740.6	347.3	912.0	2934.0	846.8	109.2	386.3	962.0
22	320	1705.4	682.8	1615.0	2405.0	566.6	182.2	379.0	786.0
23	219	2540.5	1332.6	1901.0	3235.5	1637.2	819.7	1094.1	1519.8
31	1129	1635.0	396.9	1113.0	2034.0	1153.4	361.0	716.6	1204.7
32	6829	3478.5	344.6	976.0	2689.3	2261.6	423.4	1560.5	3021.5
33	34057	10381.5	1318.9	4880.1	19502.0	4716.9	909.0	3150.0	6845.0
42	7191	12705.2	3307.3	12883.5	21303.0	4529.5	1373.3	4822.4	7585.3
44	3717	6381.8	1148.0	5751.2	11049.0	4851.3	1333.2	3250.0	8314.0
45	10105	15245.5	2408.8	6178.0	17684.0	9765.5	2656.0	5200.0	14682.0
48	269	1924.9	1353.0	1897.0	2455.0	312.0	164.0	257.0	458.0
49	17	7214.9	6957.0	7402.0	7937.0	408.6	389.0	413.0	440.0
51	3097	3919.9	1066.1	2378.2	7105.0	1053.5	0.0	381.0	1464.6
53	56	343.6	13.7	122.0	763.5	18.2	2.7	9.0	30.4
54	989	4670.7	1090.5	2103.0	10856.0	2910.2	816.5	2483.6	4858.0
56	43	1636.0	1689.0	1906.0	2044.0	84.0	75.0	80.0	92.0
61	9	358.5	342.4	358.0	366.8	28.3	27.6	29.3	31.5
62	146	488.3	63.9	573.4	785.9	48.2	4.6	54.9	77.9
71	1	14.1	14.1	14.1	14.1	3.3	3.3	3.3	3.3
72	205	1736.8	422.7	1861.0	2801.7	84.9	33.1	85.7	115.1
81	12	401.6	384.8	402.2	421.4	31.6	30.9	31.4	33.7
99	1243	11938.2	8956.0	11880.0	15133.5	8374.2	5026.0	6735.0	11744.0

6.3.1 *Single Echelon Model*

Let q_t be the order placed at time t. We consider a single-item adaptive base-stock control policy of the form:

$$q_t = d_t + L(F_{t+1} - F_t), \tag{6.1}$$

where L is the lead time, and F_{t+1} the forecast of d_t calculated at the end of period t. We assume that we (1) observe demand, (2) calculate the forecast for the next period, (3) place the orders for the period, (4) receive the orders placed L periods ago, and finally (5) fulfill demand from inventory (backlogging any demand that cannot be met). Graves (1999) adopts this policy for a multi-echelon environment with non-stationary demand, arguing that it is a myopic policy that, while not necessarily optimal, is a "reasonable extension of the order-up-to policy". The policy contains two terms; the left term replenishes the period's demand, and the right term modifies the base-stock level in proportion to the change in forecast, to account for changes in the mean lead-time demand.

We propose, for later econometric estimation, a firm-level model based upon this replenishment rule. For the purpose, we assume that a firm controls the entirety of their purchases equally and that lead times are constant. Using this model with firm-level data is in line with Rumyantsev and Netessine's (2007)[7] empirical finding that aggregate inventory changes are positively associated with sales surprise, a term introduced by Gaur et al. (2005). Sales surprise is defined as the ratio of sales to demand forecast, and thus quantifies unexpectedly high (or low) sales. Additionally, suppose that a rational decision maker on behalf of firm i uses a variant of the proposed policy whereupon she can arbitrarily adjust the order quantity by an amount u in every period:

$$q_{i,t} = d_{i,t} + L_i(F_{i,t+1} - F_{i,t}) + u_{i,t}. \tag{6.2}$$

Here, we see that the base-stock level is adjusted proportionally to the change in mean lead-time demand (as in Equation (6.1)), but is also adjusted discretionarily in every period through the adjustment quantity u_t, of which we make no assumptions yet. We call the first kind of adjustment the 'planned change in inventory buffer', and the second the 'unplanned change in inventory buffer'. For notational convenience, we denote the change in the forecast as $\Delta F_{i,t}$,

$$\Delta F_{i,t} = F_{i,t+1} - F_{i,t}, \tag{6.3}$$

then we have:

$$\underbrace{q_{i,t}}_{\text{Orders}} = \underbrace{d_{i,t}}_{\substack{\text{Demand}\\ \text{Replacement}}} + \underbrace{\beta^1 \, \Delta F_{i,t}}_{\substack{\text{Planned change}\\ \text{in inventory buffer}}} + \underbrace{u_{i,t}}_{\substack{\text{Unplanned change}\\ \text{in inventory buffer}}}, \tag{6.4}$$

where β^1 is a coefficient, to be estimated, that contains information on the replenishment lead times. When $u_{i,t} = 0$, this is equivalent to the policy from Equation (6.1) with $\beta^1 = L$.

6.3.2 *Hypotheses Development*

Changes in base-stock levels (inventory buffers), whether "planned" or "unplanned", materialize through changes in order quantities. Burns and Sivazlian (1978) show that the transient nature of such changes in downstream orders contribute to the amplification of order variability in serial supply chains because they are interpreted as persistent demand changes by upstream firms, which triggers an overreaction. Analytical work shows that information sharing can be used to reduce the amplification—with the caveat of specific demand

[7] Rumyantsev and Netessine (2007) tested several hypothesis derived from traditional inventory theory using a multiplicative inventory model with which they quantified the influence of several independent variables. Our specification differs from theirs in that we test a linear inventory model and consider both positive as well as negative deviations from the forecasted values.

distributions, or assumptions on the knowledge of the underlying replenishment rules (Chen and Lee, 2009).

Experimental work based upon behavioral operations theory, on the other hand, drops such assumptions in favor of human beings making decisions in the context of a supply chain simulation—usually some version of the beer distribution game. This stream of literature suggests that the human biases that are a cause of the Bullwhip Effect are robust to information sharing. They show that, even when controlling for all operational causes of the Bullwhip Effect, and under full information sharing, the amplification of orders in a supply chain persists (Croson and Donohue, 2006).

Our first hypothesis comes from the combination of our inventory model with the insights mentioned above. When analyzing the customer-supplier data pairs, we expect that when a downstream firm executes a planned change in their inventory buffer, the upstream firm will overreact to this change and will adapt his inventory buffer proportionally to the downstream change.

H 6.1 *Changes in upstream inventories are positively associated with planned changes in downstream inventory buffers.*

Similarly, we expect that unplanned changes in downstream inventory buffers will have a similar effect on upstream inventories.

H 6.2 *Changes in upstream inventories are positively associated with unplanned changes in downstream inventory buffers.*

6.3.3 Two-Echelon Model

To test Hypotheses 6.1 and 6.2, we develop a two-echelon model. We identify customers with the subscript c, and suppliers with the subscript s so that we can express the orders of customer c at time t as:

$$q_{c,t} = d_{c,t} + \beta^1 \Delta F_{c,t} + u_{c,t}. \tag{6.5}$$

Furthermore, we propose the following model of supplier s's orders to test hypotheses 6.1 and 6.2:

$$q_{s,t} = d_{s,t} + \beta^2 \, \Delta F_{s,t} + \beta^3 \, \beta^1 \, \Delta F_{c,t} + \beta^4 \, u_{c,t} + u_{s,t}. \tag{6.6}$$

Coefficients β^3 and β^4 respectively quantify how changes in the customer's planned, and unplanned, inventory buffers influence the supplier's ordering decisions. Because the information pertaining to the customer's changes in planned and unplanned inventory is already contained in the supplier's demand, these coefficients represent an explicit over-reaction to such inventory changes. The null hypothesis, that suppliers only respond to downstream inventory changes through the demand information, implies $\beta^3 = \beta^4 = 0$.

6.3.4 *Demand Forecasts*

In view of the fact that management's forecasts are not publicly available, we must estimate sales forecasts for every firm in the sample. This poses a challenge because we do not have information regarding the forecasting methods used in practice by the firms, and must thus make a simplifying assumption.

There is a vast body of research that attempts to quantify the performance of different forecasting methods and the subsequent decision on which method to use in different contexts. One well established method to quantify forecast performance is through the so-called forecast competitions: Empirical studies where different forecasting methods are used to forecast a large number of data series in an effort to obtain objective measures of relative performance. Makridakis and Hibon (2000) present the results of one such competition and compare the results with prior competitions. They show that more complex methods do not necessarily exhibit better performance, and that performance itself is dependent of the particular performance measure used. More importantly, they show that these results are consistent across studies.

Taking this into account, as well as the prevalence of seasonal, trending, time series in our data, we compute forecasts for every time series in our sample using the simplest seasonal-forecast method: the additive seasonal Holt-Winters forecasting procedure (Chatfield, 1978). Dekker et al. (2004), in a study aimed at developing methods to improve seasonal forecasts, show that the Holt-Winters method (HW) performs best when seasonality patterns are deterministic, and when demand variance is reduced through aggregation. The temporal and product aggregation present in our data series, combined with the relative simplicity of the HW method motivate its adoption in this study.

The HW forecasting procedure requires three smoothing constants, one each for the level (α), trend (β), and seasonal (γ) components. Formally, given the estimates for the level ($a(t)$), linear trend ($b(t)$), and seasonality ($s(t+\tau-L)$), the τ step-ahead forecast, $\hat{x}_{t+\tau}$, is defined as:

$$\hat{x}_{t+\tau} = (a(t) + b(t)\tau)\, s\,(t+\tau-L)\,, \tag{6.7}$$

where L measures the periodicity of the seasonality ($L = 4$ in quarterly data). Given the smoothing parameters α ($0 \leq \alpha \leq 1$), β ($0 \leq \beta \leq 1$), and γ ($0 \leq \gamma \leq 1$), the updating equations are:

$$a(t) = \alpha \frac{x_t}{s(t-L)} + (1-\alpha)\left(a(t-1) + b(t-1)\right), \tag{6.8}$$

$$b(t) = \beta\left(a(t) - a(t-1)\right) + (1-\beta)b(t-1), \tag{6.9}$$

$$s(t) = \gamma\left(\frac{x_t}{a(t)}\right) + (1-\gamma)s(t-L), \tag{6.10}$$

We compute the forecast for each firm using the same smoothing parameters ($\alpha = \beta = \gamma = 0.3$)[8] .

As an additional robustness check, we performed analyses using different forecasting methods. In particular, we used the time-series data to compute the best-fitting parameters for each individual firm, as well as the best-fitting sales forecasts using simple exponential smoothing and Holt's linear trend forecasting procedures. We calculated the errors for each forecasting method, for each individual series and then performed the entire analysis using the best-fitting forecast per series. The results obtained with the latter method are comparable to the results using only the HW forecasts. Therefore, we report the results obtained using the HW method. By choosing a single method for every firm we avoid introducing unwarranted complexity, as well as potential biases into the analysis (we have no reason to believe firms use the best-available forecasting method in their operations).

6.3.5 *Econometric Specification*

We use fixed-effects OLS estimation with robust standard errors (to account for heteroscedasticity in the data) to estimate the following regression model for customers:

$$\Delta i_{c,t} = a_c + \beta^1\, \Delta F_{c,t} + u_{c,t}, \tag{6.11}$$

where, from the inventory balance equation, we define $\Delta i_{c,t}$ as,

$$\Delta i_{c,t} = i_{c,t} - i_{c,t-1} = q_{c,t} - d_{c,t}. \tag{6.12}$$

[8]We settled on these values after performing an exploratory search on randomly chosen series. We found forecast performance to be quite robust within reasonable parameter values, and $\alpha = \beta = \gamma = 0.3$ to be a good compromise between the nervousness and tracking ability of the forecasts.

Here, a_c is the time-invariant firm-specific fixed effect, β^1 is the coefficient of $\Delta F_{c,t}$, and u_c, t, the unplanned change in inventory buffer, is defined as the error term for the observation. We perform the estimation on the change of inventory level, rather than on the estimated purchase quantity to reduce the effects of scale.

To test hypotheses 6.1 and 6.2, we perform a similar regression on the upstream data. For each customer-supplier pair, c-s, we have:

$$\Delta i_{s,t} = a_s + \beta^2 \, \Delta F_{s,t} + r_{c,s} \, \beta^3 \, \beta^1 \, \Delta F_{c,t} + r_{c,s} \, \beta^4 \, u_{c,t} + r_{c,s} + u_{s,t}, \tag{6.13}$$

where $r_{c,s}$ is a moderating variable derived from the relative sizes of the firms, computed through:

$$r_{c,s} = \frac{1}{T} \frac{\sum_1^T d_{s,t}}{\sum_1^T d_{c,t}}, \tag{6.14}$$

with T the length of the link between firms. We include this interaction effect because, by definition, the orders of customer c consist of an unknown fraction of the demand of s. By re-scaling the customer parameters thus, we obtain a more meaningful estimation of the regression coefficients β^3 and β^4. Additionally, we include the coefficient $r_{c,s}$ in the regression to test for any information contained in the scaling coefficient by itself.

6.4 *Safety Stock Model*

In this section, we study the unplanned change in inventory buffers in greater detail. We develop a series of hypotheses to study the relationship between the changes in inventory buffers and the economic and financial conjuncture. To test these hypotheses, we adopt the econometric structural modeling framework developed by Olivares et al. (2008) with which we impute the cost parameters of a rational newsvendor-type model to the empirical observations.

Newsvendor-style equations are common in the inventory-modeling literature. They provide a way of quantifying the trade-off between the holding- and penalty-costs that result from over- and under-ordering in a stochastic setting. While the newsvendor formulation depends on strong assumptions (single-item, single-period systems with zero lead time), studies have found that newsvendor-style equations also hold in more complex inventory systems. For example, Rogers and Tsubakitani (1991) prove that a newsvendor-type result minimizes costs in a divergent two-echelon, periodic-review inventory system with positive lead times and budgetary constrains. Further, Diks and De Kok (1998) show that in a divergent *N-echelon* system, applying nesvendor-type equations at every end-stockpoint minimizes long-run costs. In this section, we assume that the determination of the empirical order quantities in our sample follows a rational newsvendor-style model where the critical fractile (determined by the relative cost of over- and under-ordering) varies, period-by-period and firm-by-firm.

In this model, we explicitly link the unplanned changes in the inventory buffers to a cost function (represented by the critical fractile) unknown to us, but known to the decision-maker. In this view, the unplanned change in inventory buffer reflects a hedge made by the decision-maker based upon a cost structure that is no longer assumed constant, but that changes in time. Since this cost information is a priori unknown to us, to estimate it we assume that the decision-maker is rational, and that the decisions he makes, which we can observe in our data, are optimal in the context of this newsvendor-type cost model.

We observe these newsvendor decisions in the form of variable inventory buffers. In the preceding section, we introduced a replenishment model that quantified safety stock changes through two parameters: planned and unplanned changes in the inventory buffers. The former explicitly describes the variation due to a shift in the mean demand (as measured by a change in the firms's forecasts), while the latter acts as an umbrella factor that represents any other possible changes. With our newsvendor-type model, we assume that unplanned changes are driven by changes in the cost factors. After we obtain an estimate of the costs that drive the decisions, we test a series of hypothesis related to the external factors driving these costs.

6.4.1 *Newsvendor Model*

Let $C^u_{s,t}$ and $C^o_{s,t}$ represent supplier s's underage and overage costs, respectively, at time t. In the context of our model, because we assume that firms carry safety stocks, we relate these costs to the cost of increasing (or decreasing) the unplanned component of the inventory buffers. This is represented by a cost function dependent on the deviation of orders from the demand plus planned changes in the inventory buffer:

$$C^{UB}(q_{s,t}) = \left[C^o_{s,t}(q_{s,t} - (d_{s,t} + p_{s,t}))^+ + C^u_{s,t}((d_{s,t} + p_{s,t}) - q_{s,t})^+\right], \qquad (6.15)$$

where $p_{s,t}$ is the planned change in inventory buffer, and $C^{UB}(q_{s,t})$ the total cost related to changes to the unplanned inventory buffer. Here, we are not interested in assigning a defined interpretation to these costs, rather, we are interested in the relative cost ratio:

$$\gamma_{s,t} = \frac{C^o_{s,t}}{C^u_{s,t}}. \tag{6.16}$$

This cost ratio can vary within a single supplier over time, and it depends on information known by the decision-maker. When $\gamma_{s,t}$ increases, the firm has an incentive to reduce inventories; when $\gamma_{s,t}$ decreases, the firm has an incentive to increase inventories. Assuming continuous variables, the first order condition that follows from a one-period, myopic minimization of the cost Equation 6.15 results in the following condition for the optimal decision $q^*_{s,t}$:

$$F(q^*_{s,t}) = Pr(d_{s,t} + p_{s,t} \leq q^*_{s,t}) = \frac{1}{1+\gamma_{s,t}}. \tag{6.17}$$

where $d_{s,t} + p_{s,t}$ is the estimated order quantity before considering the cost components and can be estimated from our data through:

$$d_{s,t} + p_{s,t} = a_s + \beta^2\,\Delta F_{s,t} + r_{c,s}\,\beta^3\,\beta^1\,\Delta F_{c,t} + r_{c,s}\,\beta^4\,u_{c,t} + r_{c,s} + \epsilon_{s,t}, \tag{6.18}$$

$$d_{s,t} + p_{s,t} = \Omega X_{s,t} + \epsilon_{s,t}, \tag{6.19}$$

where $\epsilon_{s,t}$ represents the error in the estimation and we use Ω and $X_{s,t}$, vectors of coefficients and covariates, to represent the estimation equation in compact form.

6.4.2 Hypothesis Development

The reasoning behind the use of a policy through which the decision-maker of a firm is allowed to adjust purchase orders every period, is that there are conditions not captured by the stationary-cost model that, in practice, impact daily decision making. Examples of such conditions are batch-discount pricing, promotions (both at the supplier and customer level), advance information not captured in the forecast, the firms' financial standing, and macro economic conditions. While we cannot model the full extent of information available to a decision maker, we can track certain financial indicators and the overall macro-economic conjuncture. Consequently, we present two hypotheses that link these indicators to the variable cost ratio presented in the previous section.

Our first hypothesis comes from the combination of results we obtained in Chapter 3 and knowledge from prior empirical research. In Chapter 3, we showed that a sharp decrease in target inventories, as a reaction to the onset of the recent financial crisis, is consistent with the dynamics observed by manufacturers in the periods following Lehman Brothers' bankruptcy. In spite of the unusual magnitude and synchronization of the recent financial crisis, the mechanism proposed (that firms reduce target inventory levels when they face financial difficulties) is not new. Escaith et al. (2010), for example, suggest that reducing inventories is a common first reaction in the face of adverse credit conditions and activity slowdown. More recently, Pesch and Hoberg (2013) analyze firms facing financial distress during the 1995-2007 period and estimate that approximately 70% reduce their inventories in order to free up cash and prevent bankruptcy.

Thus, we expect to see a negative relationship between financial constraints and the cost ratio. Given this, we formulate our cost Hypotheses:

H 6.3 *When firms are constrained by liquidity, they adjust their critical fractile down.*

Similarly, we hypothesize that the overall economic sentiment will have a positive association with the acceptable inventory-related risks that a firm is willing to face; in recessionary times, firms will prioritize cash over service levels. Using the relative change in Gross Domestic Product (GDP) as a proxy to the overall economic conjuncture, we expect that negative GDP changes will shift the cost ratio balance towards penalizing overage. Therefore, we hypothesize that:

H 6.4 *When macro-economic conditions are adverse (GDP decreases), firms adjust their critical fractile down.*

6.4.3 Structural Estimation of Cost Ratios

In the standard normative approach to newsvendor-type models, the researcher assumes a distribution function for the random variable and certain explicit cost parameters. With these assumptions, he then computes the optimal ordering decision $q*$. Unfortunately, we cannot directly observe the evolution of the cost ratios for the firms in our dataset. Thus, we use a structural modeling framework to estimate the cost parameters given an assumed distribution and observations of the realized decisions.
We assume that the decision-maker is locally rational and therefore the observed orders are optimal. We use this assumption to impute, for each observation, the cost ratio that would make the observed decision rational.

Additionally, we cannot assume that two firms will react equally given a certain cost ratio. There is an ex ante heterogeneity in the inventory buffers; different decision-makers use different processes to select the target buffer inventory (e.g., at an aggregate level the inventory buffer may be associated to the number of Stock Keeping Units (SKU), the time of year, or other firm-specific factors). To model this heterogeneity in the inventory buffers, we follow Olivares et al. (2008). We assume that the value of the inventory buffer is given by independent random variables from a common family of distributions, $(F(\cdot;\theta) : \theta \in \Theta)$, where θ is a vector parameter from the parameter space Θ which characterizes each member of the class. The distribution of the inventory buffer of supplier s at time t is given by $F(\cdot;\theta_{s,t})$. We let this distribution depend on the vector of covariates $X_{s,t}$ as defined in Equation (6.18), and assume the functional form:

$$\theta_i = h(X_{s,t}, \eta), \tag{6.20}$$

where η is a vector of parameters to be estimated. Thus, the distribution of the desired inventory buffer for firm s at time t, $F(\cdot;\theta_{s,t})$ is characterized by the functional form of the distribution, the function $h(\cdot;\cdot)$, the vector η, and the vector of covariates $X_{s,t}$. In other words, all else being equal, different decision makers will have different target inventory levels in different periods.

In addition, recall that the decision-maker may face different trade-offs between overage and underage costs across observations: the cost ratio γ may differ across observations, and it depends on a series of factors that are a priori unknown to us. Formally, we define:

$$\gamma_{s,t} = g(Z_{s,t}, \alpha), \tag{6.21}$$

where Z_i is a vector of covariates, $g(\cdot)$ is a link function, and α is a vector of parameters to be estimated.

We can express the critical fracile as:

$$F(q_i^*; h(X_{s,t}, \eta)) = \frac{1}{1 + g(Z_i, \alpha)}, \tag{6.22}$$

where q_i^* specifies the optimal decision for each observation.

Following Olivares et al.'s (2008) N1 model of the decision-maker behavior, we assume that there are unobservable factors that affect the calculation of the overage/underage ratio. Let $\xi_{s,t}$ be an i.i.d (unobservable) factor that affects the calculation of cost ratio, $E(\xi_{s,t} = 0)$ and let the cost ratio follow the following log-linear specification:

$$\text{Log}\,(\gamma_{s,t}) = \alpha Z_t + \xi_t. \tag{6.23}$$

Since we do not know the realizations of $\gamma_{s,t}$, we cannot estimate α. We know, however, that if the decision-maker is rational, then she will behave according to the critical fractile:

$$\gamma_{s,t} = \frac{1}{F(q_{s,t};\theta_{s,t})} - 1 \tag{6.24}$$

The procedure to estimate α is then:

> *Step 1:* Estimate η through maximum likelihood using the observed realizations of $q_{s,t}$. Use $\hat{\eta}_{s,t}$ to compute the fitted values $\hat{\theta}_{s,t} = h(X_{s,t}, \hat{\eta})$.
> *Step 2:* Compute the estimated cost ratios $\hat{\gamma}_{s,t} = 1/F(q_{s,t};\hat{\theta}_{s,t}) - 1$, and then estimate α through an OLS estimation of Equation (6.23).
> (Olivares et al., 2008).

We assume that the distribution of the actual orders placed by the supplier at time t (demand plus planned changes in inventory buffer plus unplanned changes in inventory buffer) is defined by the vector $\theta_{s,t} = (\mu_{s,t}, \sigma_{s,t})$ and the orders placed can be written as:

$$q_{s,t} = \Omega X_{s,t} + \epsilon_{s,t}. \tag{6.25}$$

(Olivares et al., 2008) show that, if we assume $\epsilon_{s,t}$ to be i.i.d, normally distributed with mean zero and standard deviation σ, then $\eta = (\Omega, \sigma^2)$, $h(X_{s,t}, \Omega, \sigma^2) = (\Omega X_{s,t}, \sigma^2)$, and estimating (Ω, σ^2) via maximum likelihood is equivalent to estimating Ω through OLS and σ through the standard deviation of the regression residuals. Furthermore, we can estimate the critical fractile through:

$$F(d_{s,t} + p_{s,t}, \hat{\theta}_{s,t}) = \Pr\left(d_{s,t} + p_{s,t} \leq q^*_{s,t}\right) \tag{6.26}$$

$$= \Phi\left(\frac{q_{s,t} - \hat{\Omega} X_{s,t}}{\hat{\sigma}}\right). \tag{6.27}$$

And finally estimate $\gamma_{s,t}$ through:

$$\hat{\gamma}_{s,t} = \frac{1}{F(d_{s,t} + p_{s,t}, \hat{\theta}_{s,t})} - 1 \tag{6.28}$$

Since the normality assumption is quite restrictive, we repeated the analysis using the empirical distribution of the regression errors for each industry segment to calculate the critical fractile. Results obtained using said method are consistent with the results derived with this assumption.

6.4.4 Econometric Specification

We now formulate the econometric specification to test Hypotheses 6.3 and6.4. After estimating the cost ratio γ, we obtain an estimate of the vector of coefficients α through:

$$Log(\gamma_{t,s}) = \lambda^1 QuickRatio_{s,t-1} + \lambda^2 \Delta GDP_{t-1} + \lambda^3 Year_{s,t} + \lambda^4 q_2 + \lambda^5 q_3 + \lambda^6 q_4 + \xi_t. \tag{6.29}$$

Here, $QuickRatio_{s,t-1}$ is the lagged quick ratio of firm s, calculated using the financial identity:

$$QuickRatio_{s,t} = \frac{CurrentAssets_{s,t} - Inventories_{s,t}}{CurrentLiabilities_{s,t}}. \tag{6.30}$$

ΔGDP data is obtained from the Bureau of Economic Analysis[9]. To control for other external effects, we add *Year*, a linear time trend dummy; and q_2–q_4, quarterly dummies.

6.5 Results

We use the "xtreg" panel data module in STATA to perform our analysis. We estimate fixed effect regressions with robust standard errors.

Table 6.4 shows the results of the estimation of the inventory model with the change in supplier's inventory as the dependent variable. Column 1 provides the results for the pooled regression; columns 2–4 provide the results for individual industry segments as defined by the 2-digit NAICS industry code. Similarly, Column 1 of Table 6.5 provides the pooled results of the cost ratio model, with $Log(\gamma_{s,t})$ as the dependent variable; columns 2–4 show the results for the individual industry segments. Both tables show the estimation results using all periods in the sample. Table 6.3 shows the detail of the segments included in each of the NAICS codes.

The data at the aggregate level are consistent with Hypothesis 6.1: Upstream inventory changes are positively correlated with planned changes in downstream inventory buffers. This implies that upstream firms react to planned downstream inventory changes. This relationship is not statistically significant for industries that belong to NAICS code 32.

The data at the aggregate level are consistent with Hypothesis 6.2: Upstream inventory changes are positively correlated with unplanned changes in down-

[9] http://bea.gov/national/index.htm

stream inventory buffers. This implies that upstream firms react to unplanned downstream inventory changes. This relationship is statistically significant for all individual industries.

The data at the aggregate level are consistent with Hypothesis 6.3: The estimated cost ratio decreases with the lagged quick ratio. This relationship is also not significant for industry code 32.

The data at the aggregate level are consistent with Hypothesis 6.4: The estimated cost ratio decreases with the lagged change in GDP. However, this effect appears to be driven mainly by firms in industry code 33; this relationship is not significant for industry codes 31 and 32.

To add to the hypotheses tests, we analyze the significance of the other regression coefficients. The inventory regression for the industry code 32 shows that, in addition to the scaled change in downstream buffer, the firm's own change in forecast is not statistically significant. This suggests that this segment does not plan its production runs according to forecasts. A plausible explanation for this observation is that code 32 consists majorly of process industries.

The statistical significance of the dummy variables in the cost ratio regressions are consistent across all industry segments. The 3rd and 4th quarter dummies are statistically significant and positive for all industries; the 2nd quarter dummy is not statistically significant for code 31. Furthermore, the coefficients increase from q_2 to q_4, this reflects a negative association between inventory buffers and the fiscal year cycle—with buffers decreasing until they reach a minimum in the last quarter. The main reason for including the yearly trend in the cost ratio regression was to account for the potential influence that expanding product lines can have in the inventory buffers. However, the yearly trend is statistically significant only in code 32 industries, and with a positive sign. This suggests an increasing trend in the cost ratio, equivalent to a decreasing trend in buffers. Together with the results obtained through the inventory regression results for code 32 industries, these observations merit further analysis.

Table 6.3 *Segments per industry code*

Code	Industrial Segments
NAICS 31	Food, Beverage and Tobacco, Textile, Apparel, and Leather Manufacturing.
NAICS 32	Wood Product, Paper, Printing, Petroleum and Coal, Chemical, Plastics and Rubber, and Non-metallic Mineral Product Manufacturing
NAICS 33	Primary Metal, Fabricated Metal, Machinery, Computer and Electronic Product, Electrical Equipment, Appliance and Component, Transportation Equipment, Furniture, and Miscellaneous Manufacturing.

Table 6.4 *Pooled and industry-specific inventory regressions for period 1984-2013*

Coefficient	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔF_t	0.055	(0.020)***	0.182	(0.037)***	0.005	(0.025)	0.071	(0.023)***
Scale	0.153	(0.765)	7.949	(4.607)*	-1.210	(1.306)	0.304	(0.530)
Scaled Unplanned Buffer	0.120	(0.032)***	0.206	(0.099)**	0.226	(0.079)***	0.097	(0.033)***
Scaled Planned Buffer	0.446	(0.135)***	0.680	(0.316)**	0.179	(0.191)	0.437	(0.147)***
Constant	1.760	(0.131)***	0.578	(0.344)*	4.059	(0.165)***	1.110	(0.117)***
N	71337		7463		15662		48212	
Suppliers	2558		275		624		1659	

Table 6.5 *Pooled and industry-specific cost regressions for period 1984-2013*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔGDP_{t-1}	-0.011	(0.003)***	0.006	(0.005)	0.006	(0.011)	-0.020	(0.003)***
$QuickRatio_{t-1}$	-0.005	(0.001)***	-0.065	(0.015)***	-0.001	(0.001)	-0.008	(0.001)***
Scaled Unplanned Buffer	0.001	(0.001)	-0.003	(0.003)	0.000	(0.003)	0.001	(0.001)
Scaled Planned Buffer	-0.004	(0.003)	-0.005	(0.008)	0.002	(0.004)	-0.004	(0.003)
Year	0.001	(0.003)	0.005	(0.005)	0.013	(0.006)**	-0.003	(0.004)
q_2	0.047	(0.015)***	-0.046	(0.059)	0.133	(0.038)***	0.032	(0.016)**
q_3	0.161	(0.019)***	0.360	(0.091)***	0.224	(0.053)***	0.115	(0.019)***
q_4	0.274	(0.026)***	0.399	(0.106)***	0.286	(0.066)***	0.240	(0.026)***
Constant	-0.059	(0.067)	-0.181	(0.142)	-0.404	(0.148)***	0.061	(0.080)
N	70,962		7,441		15,617		47,904	

Note: Standard errors are reported in parentheses.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

6.5.1 *The Impact of the 2008 Financial Collapse*

The recent credit crisis was global in nature, dramatic in magnitude, and significantly affected the performance of manufacturing firms (Levchenko et al., 2010). In Chapter 3, we argue that pressing financial conditions drove individual firms to seek monetary refuge by converting inventories into cash, and find support support for our hypothesis at the individual supply chain level. Similar observations, concerning higher aggregation levels, have been made in the financial literature: Gao and Yun (2009), for example, explicitly link manufacturing performance in the periods following Lehman brothers' bankruptcy to firms' ability to withstand the financial turmoil. They show that firms' access to liquidity is positively associated with their business performance during the 2008-2009 period.

The results we have presented thus far in this chapter take explicit consideration of financial conditions. To account for the potential impact that significant[10] crisis - related disruptions in our results, we temporally disaggregate our dataset

[10]As in behavior-changing paradigm shifts.

and repeat our econometric analysis independently for the 1984–2006 and 2007–2013 periods.

Tables 6.6 and 6.7 provide the results of the inventory regressions for the periods 1984–2006 (pre-crisis) and 2007–2013 (crisis/post-crisis) respectively. Similarly, Tables 6.8 and 6.9 show the results of the cost ratio regressions for those periods.

The temporally disaggregated data are consistent with the findings detailed in the previous section. All hypotheses are supported by the pooled data.

At the industrial segment level, it is interesting to note that the results from the pre- and post-crisis inventory regressions are generally consistent for codes 32 and 33, but not for code 31. We see, in the pre-crisis period (column 2 of Table 6.6), no statistical significance for either of the coefficients related to downstream inventory buffers. This changes significantly in column 2 of Table 6.7.

To test for a change in the behavior of the firms among the periods, we perform a statistical test of the values of the coefficients for pre- and post-crisis regressions. We do so by generating a dummy variable to indicate whether an observation belongs to the pre- or post-crisis period and then re-running the regression with an interaction term between this indicator variable and each of the predictors. Then, the *p-value* for the interaction term effectively gives us a significance test for the difference between the coefficients. We find that, at the aggregate level, the difference between the pre- and post-crisis Scaled Planned Buffer coefficients is statistically significant at the 10% level. None of the other coefficients show a statistical difference across periods.

Comparing the results of the cost ratio regressions, results at the aggregate and industry segment levels are generally consistent. A notable observation is that the quick ratio coefficient for code 32, which was not statistically significant in the previous analysis becomes significant and negative in the post-crisis regression (column 3, Table 6.9). This suggests that process industries also steered on cash during and after the credit crisis. Also notable is the statistical significance and negative sign of the yearly trend dummy in the pre-crisis regression. This is mainly driven by code 33 (column 3, Table 6.8) and suggests a progressive increase in the desired inventory buffers through this period. This relationship turns statistically non-significant in the post crisis period. However, further research is needed to identify the causality of the '84-'06 relationship (an increase in the number of SKU's is a plausible explanation), and the reasons why this relationship is not observed in the post-crisis period.

Finally, we perform statistical tests to quantify the change in the coefficients. We find that the quick ratio coefficient during the post-crisis period is more negative than during the pre-crisis period with a statistical significance at the 1% level; the quarterly dummies for 3rd and 4th quarter, on the other hand, are found to increase in the post crisis period with the same level of statistical significance. This underscores the increased importance of the financial performance during the period inasmuch as it highlights an increase in the association between liquidity and inventory buffers, and between the fiscal year and inventory buffers.

Table 6.6 *Estimation results for period 1984-2006*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔF_t	0.0460	(0.0179)**	0.146	(0.0491)***	0.0644	(0.0489)	0.0395	(0.0185)**
Scale	0.0291	(0.742)	0.612	(2.104)	0.731	(0.432)*	0.0200	(0.804)
Scaled Unplanned Buffer	0.0843	(0.0420)**	0.124	(0.0935)	0.0590	(0.0278)**	0.0837	(0.0461)*
Scaled Planned Buffer	0.597	(0.169)***	0.215	(0.407)	0.236	(0.170)	0.627	(0.182)***
Constant	1.029	(0.126)***	-0.0728	(0.146)	2.268	(0.075)***	0.871	(0.151)***
N	51029		5404		9758		35867	
Suppliers	2235		249		502		1484	

Table 6.7 *Estimation results for period 2007-2013*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔF_t	0.0602	(0.0328)*	0.242	(0.0420)***	-0.000183	(0.0229)	0.130	(0.0450)***
Scale	0.292	(1.702)	11.623	(6.684)*	-0.0439	(1.953)	1.613	(0.682)**
Scaled Unplanned Buffer	0.187	(0.0555)***	0.266	(0.113)**	0.341	(0.131)**	0.117	(0.0440)***
Scaled Planned Buffer	0.259	(0.131)**	0.735	(0.300)**	-0.131	(0.354)	0.232	(0.135)*
Constant	3.593	(0.306)***	3.025	(0.677)***	6.517	(0.268)***	1.826	(0.100)***
N	20308		2059		5904		12345	
Suppliers	906		99		274		533	

Note: Standard errors are reported in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 6.8 *Pooled and segment-specific cost regressions for period1984-2006*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔGDP_{t-1}	-0.007	(0.002)***	0.011	(0.005)**	0.004	(0.003)	-0.014	(0.003)***
$QuickRatio_{t-1}$	-0.003	(0.001)***	-0.041	(0.010)***	0.000	(0.000)	-0.006	(0.001)***
Scaled Unplanned Buffer	0.001	(0.001)	-0.002	(0.002)	-0.001	(0.002)	0.001	(0.001)
Scaled Planned Buffer	-0.001	(0.003)	0.014	(0.009)	0.003	(0.003)	-0.002	(0.003)
Year	-0.004	(0.002)**	-0.000	(0.002)	0.001	(0.003)	-0.006	(0.002)***
q2	0.041	(0.014)***	0.017	(0.042)	0.080	(0.030)***	0.028	(0.016)*
q3	0.106	(0.016)***	0.234	(0.055)***	0.121	(0.038)***	0.080	(0.019)***
q4	0.107	(0.020)***	0.254	(0.055)***	0.134	(0.046)***	0.159	(0.022)***
Constant	0.038	(0.030)	-0.063	(0.054)	-0.133	(0.059)**	0.107	(0.038)***
N	50,745		5,391		9,745		35,609	

Table 6.9 *Pooled and segment-specific cost regressions for period 2007-2013*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔGDP_{t-1}	-0.018	(0.007)***	-0.019	(0.011)*	0.009	(0.021)	-0.032	(0.006)***
$QuickRatio_{t-1}$	-0.022	(0.004)***	-0.258	(0.078)***	-0.027	(0.008)***	-0.016	(0.004)***
Scaled Unplanned Buffer	0.000	(0.001)	-0.004	(0.003)	0.002	(0.003)	0.000	(0.002)
Scaled Planned Buffer	-0.005	(0.006)	-0.007	(0.008)	-0.002	(0.006)	-0.005	(0.007)
Year	0.006	(0.010)	0.055	(0.039)	0.012	(0.024)	-0.006	(0.011)
q2	0.036	(0.037)	-0.272	(0.164)*	0.212	(0.075)***	0.014	(0.042)
q3	0.289	(0.046)***	0.632	(0.243)**	0.371	(0.099)***	0.202	(0.043)***
q4	0.515	(0.056)***	0.812	(0.287)***	0.515	(0.106)***	0.443	(0.059)***
Constant	-0.228	(0.298)	-1.503	(1.193)	-0.386	(0.682)	0.120	(0.312)
N	20,217		2,050		5,872		12,295	

Note: Standard errors are reported in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

6.6 Conclusions

In this Chapter, we have studied the effect of changes of downstream inventories in the production decisions made upstream. We find support for the hypothesis that planned and unplanned changes in a customer's inventory trigger an overreaction on the part of its supplier. Furthermore, we also find support the hypothesis that decision makers adjust their safety stock (and thus the unplanned component of their inventory changes) according to the economic sentiment (measured by the change in the GDP) and their financial position (measured as the quick ratio).

To study the relationship between downstream and upstream companies, we constructed a database of customer/supplier pairs by taking advantage of a US regulation in which customers that account for at least 10% of a firms' sales must be included in financial reports. This kind of database is not, however, in widespread use because of the format in which the customer names are reported: rather than reporting a unique and predefined identifier, firms report their customers' names in the form of a plain text string. This causes difficulties for the researcher constructing links from this information, for the data is plagued with spelling mistakes, abbreviations, and unconventional spellings. We overcome this ambiguity in the reporting using a combination of a partial string matching algorithm and manual matching. Our database consists of financial information of supplier-customer pairs during approximately 80000 firm-quarters, in the period 1984-2013.

To test our first set of hypotheses, we derive an econometric specification model that allows us to separate between the customer's planned and unplanned inventory changes by assuming that planned inventory changes follow the firms' forecast and the unplanned changes depend on factors that we are unable to observe. We find that suppliers consistently overreact to the inventory changes of their customers. This suggests that inventories are being used systematically as drivers of ordering decisions.
The second part of our study expands on this finding with a structural model of the decision making mechanism. In this model, the decision maker is assumed to be following a rational, newsvendor-like policy, and adjusting the safety stock component every period by allowing the cost ratio (the ratio between the overage and underage cost) vary. With this model, we find support for the second set of hypotheses. Moreover, and in line with the findings from chapter 3, we show that the relative importance of the firm's liquidity as a driver increased in the period 2007-2013 as compared to the 1984-2006 period.

Our study has several limitations. The use of aggregated financial data from Compustat can lead to space and time aggregation biases. We also use proxies that may introduce biases: sales for demand and production for orders. In the structural estimation, we assume a rational decision maker, which assumes knowledge and application of optimal policies. The construction of the supplier/customer pairs also brings limitations. The sample is biased towards large customers and small suppliers, which can potentially cause an over-estimation of the effects that downstream changes have upstream.

Appendix C
The Influence of firm size bias

In §6.2.1 we observed that, due to the nature of the reporting process, our customer-supplier pair database is biased towards large customers and small suppliers. This bias can potentially affect the generality of our results. It is possible that smaller firms behave differently from larger firms. Smaller firms can be more agile in implementing changes, as well as more sensitive to changes in their liquidity.

To test whether the results obtained from our analysis are being driven primarily by this bias towards smaller suppliers, we partition our data according to the relative size of customer-supplier pairs and repeat our analysis on a sub-sample of our data, comprised of the top 25th percentile. We present the results in tables C.1 through C.6.

We see that qualitatively, the results obtained with this sub-sample are consistent with those obtained through the analysis of the entire dataset. This suggests that the firm size bias is not driving our results and thus increases the confidence on the results. However, it is important to note that while this rules out the influence of firm size bias in our dataset, it is not a stringent test of the influence of firm size.

Table C.1 *Estimation results for period 1984-2013*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔF_t	0.053	(0.026)**	0.236	(0.070)***	-0.014	(0.024)	0.090	(0.029)***
Scale	-0.009	(0.895)	7.567	(4.875)	-1.957	(1.379)	0.197	(0.549)
Scaled Unplanned Buffer	0.120	(0.034)***	0.225	(0.112)**	0.241	(0.084)***	0.095	(0.034)***
Scaled Planned Buffer	0.440	(0.137)***	0.651	(0.360)*	0.133	(0.202)	0.420	(0.145)***
Constant	4.388	(0.512)***	3.577	(1.390)**	10.297	(0.515)***	2.757	(0.362)***
N	17,757		1,602		3,248		12,907	

Table C.2 *Estimation results for period 1984-2006*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔF_t	0.061	(0.029)**	0.141	(0.061)**	0.038	(0.077)	0.059	(0.030)*
Scale	0.017	(0.779)	1.118	(2.974)	0.915	(0.423)**	-0.023	(0.835)
Scaled Unplanned Buffer	0.087	(0.045)*	0.163	(0.104)	0.067	(0.027)**	0.085	(0.049)*
Scaled Planned Buffer	0.583	(0.165)***	0.023	(0.496)	0.251	(0.174)	0.610	(0.176)***
Constant	2.422	(0.449)***	0.789	(0.742)	3.584	(0.188)***	2.323	(0.545)***
N	11,727		1,023		1,980		8,724	

Table C.3 *Estimation results for period 2006-2013*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔF_t	0.045	(0.034)	0.379	(0.085)***	-0.012	(0.024)	0.123	(0.047)***
Scale	-0.292	(2.035)	9.529	(5.203)*	-0.728	(2.114)	1.573	(0.687)**
Scaled Unplanned Buffer	0.182	(0.058)***	0.277	(0.128)**	0.366	(0.143)**	0.110	(0.044)**
Scaled Planned Buffer	0.260	(0.135)*	0.708	(0.346)**	-0.247	(0.370)	0.232	(0.138)*
Constant	8.080	(1.059)***	10.124	(1.719)***	18.004	(0.813)***	3.494	(0.249)***
N	6,030		579		1,268		4,183	

Note: Standard errors are reported in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table C.4 *Pooled and segment-specific cost regressions for period 1984-2013*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔGDP_{t-1}	-0.045	(0.019)**	-0.001	(0.020)	0.005	(0.066)	-0.065	(0.020)***
$QuickRatio_{t-1}$	-0.011	(0.005)**	-0.323	(0.161)**	-0.000	(0.003)	-0.024	(0.011)**
Scaled Unplanned Buffer	0.001	(0.001)	-0.003	(0.003)	0.001	(0.003)	0.001	(0.001)
Scaled Planned Buffer	-0.000	(0.003)	-0.003	(0.012)	0.003	(0.005)	0.000	(0.003)
Year	0.010	(0.011)	0.006	(0.027)	0.050	(0.037)	0.005	(0.012)
q2	0.137	(0.075)*	-0.179	(0.259)	0.437	(0.189)**	0.115	(0.086)
q3	0.429	(0.102)***	0.744	(0.396)*	0.694	(0.285)**	0.342	(0.112)***
q4	0.782	(0.127)***	0.823	(0.500)	1.024	(0.346)***	0.696	(0.138)***
Constant	-0.292	(0.264)	-0.071	(0.943)	-1.344	(0.954)	-0.079	(0.284)
N	17,503		1,593		3,216		12,694	

Table C.5 *Pooled and segment-specific cost regressions for period 1984-2006*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔGDP_{t-1}	-0.046	(0.020)**	0.039	(0.020)*	0.011	(0.014)	-0.057	(0.026)**
$QuickRatio_{t-1}$	-0.009	(0.005)*	-0.224	(0.096)**	0.001	(0.001)	-0.017	(0.011)
Scaled Unplanned Buffer	0.001	(0.001)	-0.002	(0.002)	-0.001	(0.002)	0.001	(0.001)
Scaled Planned Buffer	0.000	(0.003)	0.017	(0.012)	0.004	(0.004)	0.001	(0.002)
Year	0.010	(0.011)	-0.014	(0.009)	0.011	(0.012)	-0.007	(0.013)
q2	0.169	(0.076)**	-0.019	(0.205)	0.231	(0.164)	0.114	(0.086)
q3	0.412	(0.102)***	0.306	(0.189)	0.353	(0.210)*	0.238	(0.113)**
q4	0.739	(0.123)***	0.406	(0.255)	0.508	(0.226)**	0.505	(0.149)***
Constant	-0.259	(0.266)	0.287	(0.310)	-0.531	(0.307)*	0.174	(0.240)
N	16,930		1,020		1,974		8,554	

Table C.6 *Pooled and segment-specific cost regressions for period 1984-2013*

Segment	Pooled		NAICS 31		NAICS 32		NAICS 33	
ΔGDP_{t-1}	-0.048	(0.020)**	-0.089	(0.034)**	0.017	(0.125)	-0.081	(0.023)***
$QuickRatio_{t-1}$	-0.010	(0.005)*	-0.623	(0.613)	-0.139	(0.112)	-0.060	(0.031)**
Scaled Unplanned Buffer	0.001	(0.001)	-0.004	(0.003)	0.003	(0.003)	0.001	(0.001)
Scaled Planned Buffer	-0.000	(0.003)	-0.003	(0.012)	-0.001	(0.007)	-0.001	(0.006)
Year	0.010	(0.011)	0.126	(0.169)	0.019	(0.162)	-0.007	(0.034)
q2	0.139	(0.078)*	-0.661	(0.517)	0.698	(0.439)	0.022	(0.165)
q3	0.433	(0.108) ***	1.447	(1.003)	1.109	(0.632)*	0.510	(0.210)**
q4	0.801	(0.130)***	1.734	(1.060)	1.716	(0.560)*	1.003	(0.216)***
Constant	-0.290	(0.276)	-3.303	(5.778)	-0.343	(4.555)	0.163	(0.985)
N	16,483		573		1,242		4,140	

Note: Standard errors are reported in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

"Would you tell me, please, which way I ought to go from here?"
"That depends a good deal on where you want to get to."
"I don't much care where –"
"Then it doesn't matter which way you go".

Lewis Carroll

Chapter 7
Conclusions

Through the years, the Bullwhip Effect has been a fertile source of research. A Google scholar search shows that, since 1997, over 1000 published scientific articles have "the Bullwhip Effect" in their title. Moreover, bullwhip research appears to be 'methodologically-agnostic': Analytical, empirical, experimental, behavioral, time domain, frequency domain, micro, macro—essentially every kind of methodology or framework has something to add to our understanding of the bullwhip.

Such a broad pursuit is driven in part because of how *interesting* the Bullwhip Effect is. The bullwhip is so encompassing that its massive reach would be a dissertation topic in itself, were it not eclipsed by the sheer *importance* it has in the operation of real life supply chains—there is still a large amount of relevant research to be conducted.

Because of this timely, practical, and scientific relevance, the main objective of this dissertation is to expand our collective knowledge of how the bullwhip and inventories are interrelated. Reading the literature review presented in Chapter 2, one cannot but notice that a discussion of the Bullwhip Effect implies a discussion of inventories and inventory policies. A fitting analogy is that inventories are the medium through which the bullwhip propagates; any study of the bullwhip is at some level a study of inventories.

One of the main contributions of this dissertation is to highlight the explicit influence that inventories have on the bullwhip, and that the bullwhip has on inventories. Methodologically, we make contributions to the simulation-based system dynamics literature (Chapter 3); the analytical, and numerical, control theory literature (Chapter 4 and Chapter 5); and to the empirical literature in operations management (Chapter 6).

The rest of this chapter is organized as follows. In Section 7.1 we summarize the findings presented throughout the thesis by answering the research questions posed in Chapter 1. In Section 7.2 we reflect upon the managerial implications that are derived from our work. We end with a discussion on future research directions in Section 7.3.

7.1 Research Questions Revisited

Since we explore the relationship between inventories and the bullwhip throughout the entire dissertation, the first research question we analyzed was:

> *What have researchers learned about the link between inventories and the Bullwhip Effect?*

In Chapter 2, we surveyed over 50 years of Bullwhip Effect research. We tackled the literature review by classifying existing research according to its main purpose: Bullwhip discovery and measurement, bullwhip analysis, and bullwhip taming. While it is immediately apparent that inventories are of extreme importance in bullwhip research, inventories are seldom the focal point of analysis. This, however, was not always so: In the original descriptions of the "Forrester Effect", before the term bullwhip was even conceived, the dynamics of inventories were as much the focus as the dynamics of orders.

Because order variability has more immediate cost (and personnel) repercussions than inventory variability, and because direct costs can more readily be imputed to highly variable orders (e.g. production cost curves are usually assumed to be steeper than inventory cost curves), order variance emerged as the key metric for the quantification of the bullwhip. In Lee et al.'s (1997a) seminal paper, the Bullwhip Effect itself is defined as the appearance of *demand distortion*. It is recognized that the distortion (amplification) originates from inventory decisions, but the measurement of inventory variability is not in itself a standard measure of the bullwhip—with the exception of work in the control theory literature, where inventory variability is explicitly taken into account when considering performance trade-offs (Dejonckheere et al., 2003; Disney et al., 2006a; Hoberg et al., 2007b).

Even though inventory variability is not the key metric of most of the bullwhip models, inventories themselves are a key part of them. Again, inventories are the medium through which the bullwhip propagates. In the vast majority of the Operations Management/Operations Research literature, inventories drive purchase decisions. Order-Up-To Policies (OUT) are implemented in statistical modeling (Lee et al., 1997a), control theoretic modeling (Dejonckheere et al., 2003), system dynamics modeling (Kim and Springer, 2008), empirical work (Bray and Mendelson, 2012), and experimental work (Sterman, 1989).

In the macroeconomics literature, on the contrary, the production smoothing model is the basis for partial and general equilibrium models that consider total inventories. In this view, inventories are not a driver anymore—they are an adjustment variable. Under the assumption of convex production costs, production smoothing justifies the usage of inventories as an adjustment variable by proposing that the optimal production schedule is a constant production schedule. Thus, inventories act as a buffer, preventing production fluctuations. As the inspection of any undergraduate or graduate texts in the field easily proves, the interest that macro-economists as a whole have on inventories is limited (for example Abel et al., 2008, contains but 3 mentions of inventories among its 750 pages). Nevertheless, a number of researchers in the field have noted that inventories can also be more than an adjustment variable.

In the macroeconomics world, the exact role of inventories is still very much an open question. Researchers in the early 1980's have pronounced the production smoothing model "in trouble" for its inability to explain the variability of orders and pro-cyclicality of inventories (traits of the bullwhip), and started advocating for the adoption of models with a micro-foundation—such as OUT replenishment policies. In the intervening years, such models have been developed, as well as modifications to the production smoothing model that make its predictions fall in line with empirical observations. These modified models smooth production, but allow decisions to depend on explicit inventory policies. As a result, inventories are both a drivers and an adjustment variable.

In OM research, the view of inventories as drivers of replenishment policies is widely accepted. There is, however, limited research on the explicit influence of inventory decisions over and above these policies. What happens if parameters change? For example, what are the effects of changes in the underlying cost structure on the supply chain dynamics? Can we infer the influence of human decision-making by studying aggregate data? The majority of the work done in this area is encompassed on the behavioral operations research, mostly within the beer game framework. This view of inventories as explicit decision-making instruments motivates the next research question.

> *Can a synchronized inventory shock –caused by the desire of firms to retain liquidity in moments of financial distress– explain the demand dynamics experienced by upstream manufacturing firms following the collapse, on*

September 2008, of Lehman Brothers?

We explored this in Chapter 3. As a testament of the importance of the Bullwhip Effect in supply chain performance, the motivation behind the development of this question came to us from the industry. In particular, this chapter is the result of the collaboration with a dutch chemical firm that was experiencing demand drops that did not correspond to what the end-markets themselves were experiencing. For example, while the sale of bottled beverages stayed stable during the first periods of the crisis, the demand of raw materials used in the manufacturing of the bottles and labels saw a significant downswing.

During the same period, higher management at this site implemented a set of crisis measures as a direct reaction to the observed plummeting of the demand. Reducing inventories across the board was one of them. Informal interviews with customers and competitors confirmed that this was the norm in the industry. We hypothesized that a synchronized reduction of inventory targets across the supply chain would result in significant demand drops upstream—the very same thing that motivated the inventory reductions in the first place.

We developed supply chain models based on system dynamics theories to test this hypothesis. From a scientific point of view, the fit of the model predictions, coupled with the failure of alternative models to explain the observations, suggests that inventory reductions played indeed a large part in supply chain dynamics during the credit crisis. The lack of data for downstream echelons of the modeled supply chains, however, lessens the power of the statistical tests that we are able to carry out.

From a practical point of view, the models we developed where incorporated into the decision system for various business units within the firm, and were used to complement medium term forecasts. These forecasts, combined with other crisis measures, enabled the firm to successfully steer through the credit crisis period.

The motivation for the next research question stems directly from the estimation of the parameters for the system dynamics models. We found that the best model fit was associated with parameters that are consistent with Sterman's (1989) beer game experiments. In particular, our results are consistent with the behavioral characteristic of slow adjustment speeds and underestimation of pipeline inventories. The next question, naturally, is:

How does human behavior -as measured by the inventory and pipeline smoothing– affect the stability of a production/inventory system, its dynamic performance, and the amplification of orders and inventories?

In Chapter 4, we contribute to the literature by presenting a compact mathematical expression for the stability of an Automatic Pipeline, Variable Inventory Order Based Production Control System (APVIOBPCS, a generalization of the OUT policy). To find this compact expression, we used linear control theory as a modeling methodology. The stability of the system[1] is completely characterized by a set of structural and behavioral parameters. In the second half of Chapter 4, we explicitly look at the effect of a particular behavior: the underestimation of the pipeline inventories:

> *How does the under-estimation of the pipeline affect the performance of the firm?*

Under-estimation of the pipeline, as we saw in Chapter 3, is a trait associated with human behavior that we observe in experimental and empirical data. Numerous studies have shown that humans playing the beer game tend to consistently under-estimate the pipeline inventory (Sterman, 1989; Croson and Donohue, 2006); it is simply difficult for us to mentally keep track of cumulative quantities. We have shown in Chapter 3 that under-estimation of the pipeline is present at firm level data, and consequently studied this phenomenon in detail in Chapter 4. We introduced stationary, transient, and steady state performance metrics to analyze the performance of the system. We identified the trade-off between bullwhip smoothing and transient performance, and found that under-estimating the pipeline can lead to marginal performance gains. However, we also observed that under-estimating the pipeline leads to peaks in the low frequency response of orders, which in turn leads to oscillations. We investigated the performance metrics and the oscillatory response through the next research question:

> *How robust are the theoretical performance metrics to changes in the demand? How does the cyclicality of the system interact with cyclical demands?*

The theoretical performance metrics we derived in Chapter 4 depend on the assumption of a demand distribution. In Chapter 5, we performed numerical experiments, with the control theoretic model, to extend our insights to demands that more accurately resemble those which we encounter in real life. Surprisingly, we found that, for certain behavioral conditions, the control theoretic models are extremely sensitive to the input demand. The way the

[1] A stable system is a system that given a finite input, gives a finite output (i.e., does not explode).

system amplifies certain frequencies more than others means that even in the presence of demand sampled from a theoretical distribution, the sampling noise is enough to bring about a response that, when measured, differs from the theoretical response. In the case of cyclical (seasonal) demands, we observed that the frequency of the demand cycles is of utmost importance. The performance associated with different behavioral parameters depends on the dominant seasonality of the demand. Therefore, demand seasonality must be explicitly considered when designing inventory control systems.

In Chapter 6, we turned to secondary data to understand more about the consequences of a number of the insights we had developed thus far. In this chapter, we looked at inventories as explicit drivers of demand dynamics. The first part of this chapter develops a series of hypothesis designed to answer the question:

> *Do upstream decision makers over-react to changes in downstream inventory levels?*

We compiled a database that contains financial information of supplier/customer pairs for the 1984-2013 period. With it, we estimated explicit changes to downstream (customer) inventory buffers, and categorized them as either planned or unplanned (the former respond to a change in the demand forecast). We found support for the hypothesis that downstream changes in inventory buffers are associated with an overreaction upstream. Our finding suggests that changes in downstream base stock policies are also transmitted upstream. This result again highlights the importance of studying inventory changes in more detail. The next, and last, question follows:

> *Can we observe empirical evidence of firms adapting their inventory levels to economic and financial conditions?*

In the second part of Chapter 6, we tested the hypothesis that economic and financial conditions constrain the target levels of inventory buffers. We have offered anecdotal evidence of such a mechanism as one of the motivations behind the study of Chapter 3, where we found support for it in the context of the credit crisis. We use a structural model of a rational decision making process to show that a cost ratio that governs a variable safety stock, dependent on the greater economic climate and on a firm's individual liquidity is consistent with the empirical observations. Additionally, we found that firms have increased their reliance on liquidity conditions during and after the crisis, which supports the assumptions of Chapter 3.

In conclusion, there are various threads that we find when we analyze the different research questions dealt with in this dissertation. First, we observe how inventories and orders are intertwined and how one affects the other no matter the modeling methodology. Second, from a practical perspective, we see that different modeling methodologies are useful in explaining different aspects of real-life observations. This has the theoretical consequence that inventory research, inasmuch as it attempts to develop insights applicable to real-life situations, must therefore embrace different methodologies to study different aspects of performance.

7.2 *Managerial Implications*

Even though the Bullwhip Effect is usually studied with particular attention to the influence of demand shocks or variations, we have seen (in chapters 3 and 6) that managerial decisions at every level of a supply chain can have a large impact, potentially generating "endogenous bullwhips". The managerial decisions that can have such an impact on performance relate to locally optimal inventory decisions; a decision that is optimal in the short term for one firm may very well be detrimental in the medium to long term, as well as when the entire supply chain's performance is considered.

A significant amount of analytical inventory models tend to assume that inventory policies are static with regards to the underlying cost structure and consequently study steady state performance. In other words, models assume that all costs remain constant relative to one another, and study the stationary consequences of a policy. This is a perfectly reasonable assumption from a theoretical point of view[2], however, we have seen evidence that suggests that in real life the cost ratios fluctuate. Viewed from a pragmatic perspective, the cost of having too much inventory is not the same to a "healthy" company than to a company with financial constraints—nor is the definition of a "healthy" company constant in time. In this thesis we have seen that, in the form of inventory adjustments, fluctuations in cost ratios can have a significant effect in the performance of a supply chain.

From a managerial perspective, this raises several interesting observations.

The first is that decisions taken in the interest of the financial short term can have repercussions in the entire performance of the supply chain. A sudden reduction of inventories will (all else equal) increase the liquidity of a firm, but may

[2]The steady state analysis of performance changes due to parameter changes is trivial and not interesting from a research perspective.

generate a Bullwhip Effect that will propagate upstream. In a capacitated supply chain, the Bullwhip Effect can cause significant stock-outs in the upswings. Thus, to downstream supply chain members, the bullwhip can return from the supply side in the form of shortages when demand picks up and over-supply when those shortages are resolved. Anecdotal evidence from the 2008 credit crisis suggests that those firms that built up stocks early profited by, for example, increasing their market share when direct competitors faced large scale stock-outs.

Second, from a supply chain perspective this implies that not all fluctuations in one's demand are created equal. This implies that the origin of, as well as the magnitude of demand fluctuations should be tracked. The knowledge of the source of demand fluctuations and shocks can lead a firm to make better informed production/ordering decisions.

In general, the main insights that can be derived from this work are summed up in the realization that not only all demand is not equal, but that the inherent dynamics of a firm mean that the idea of a target stock is necessarily a dynamic one. As to why this is important, we start by thinking about Chapter 6. The evidence presented in this chapter points to systematic adjustments of target stocks that follow economic and financial developments. In Chapter 3, we see that such dynamics are present in a plausible explanation of observations during the financial crisis. Finally, Chapter 5 shows that even in random conditions, inventories and orders fluctuate because they have an inherent dynamic signature to them. Hence, if decisions are made with the support of static data (e.g. a snapshot of inventory levels without dynamic considerations), then they can lead to amplification and oscillations (as seen in Chapter 3) which will in turn cause further adjustments.

Thus, decisions on inventories must be made taking their dynamics into consideration. Given that each individual firm possesses its own characteristics (and thus it's own dynamic footprint) and that different demand dynamics interact differently to different firms (a firm that smooths significantly will react very differently to a seasonal demand than an agile firm), it follows that both have to be fundamentally understood if correct decisions are to be made. In this, the type of decisions is as important as the decision itself. The correct reaction to a demand shock derived from a downstream inventory correction is very different to the correct reaction to a demand shock derived from a fundamental change in the downstream end-market demand. If decision-making is static, both shocks will be treated equally. If decision-making incorporates dynamics, then both shocks will be treated differently—and in some cases, no reaction will be the best reaction.

Information sharing is widely lauded as an effective way of mitigating the Bullwhip Effect. It is, at its core, a way of qualifying demand signals. In this sense, the objective of sharing information is to be able to separate between demand uncertainty and demand variability. Studies show that the majority of the benefits of information sharing can be obtained if the downstream demand as well as the inventory policy is known (Chen, 2003).

The majority of theoretical information sharing literature concentrates in the analysis of pairs of firms, or linear supply chains. We have shown, in Chapter 3, that for firms involved in more general, divergent, supply chains there is value in obtaining information regarding the source of demand changes by tracking downstream demand and inventory policies, be it by systematic study of one's own supply chain or otherwise; this view recognizes the challenge of large scale information sharing as well as its benefits. In the absence of true information sharing at these levels, then, learning about the demand patterns affecting one's own supply chain, enables firms to implement decision support systems that are capable of distinguishing between demand uncertainty and demand variability, and thus react accordingly.

7.3 Future Research

As it is not unusual in scientific research, the research contained in this dissertation raises a large number of questions. We discuss some general directions for future research concerning the main threads of this thesis through the formulation of a number of new research questions. In this section, we present a series of questions to be addressed by future research, followed by the intuition and motivation behind them.

> *Can we find further evidence of managerial underestimation of the pipeline inventory in firm level data?*

Behavioral experiments have shown time and time again that human decision-makers underestimate the pipeline inventory when playing the beer game. We have found, in chapter 3, evidence that suggests that the same effect is present in industrial settings. However, further research is needed to test this hypothesis and rule out the effects of aggregation. The main hurdle for this direction of research is the availability of data.

> *Can we learn more with better data?*

The large majority of empirical research on inventories is carried out with what are essentially single-echelon models. This is not because of a limitation in empirical modeling of supply chains, rather, it is because the availability of 'multi-echelon' data is extremely poor (as exhibited by the limitations of the data pertaining customer-supplier dyads discussed in Chapter 6). Better data, of a higher frequency and at an aggregation level that lets us construct empirical supply chains will open the doors to substantial advances in our understanding of the way that firms operate. Bray and Mendelson (2013) demonstrate the use of high-frequency (albeit single-echelon) data in a recent study of the Bullwhip Effect in the automotive industry. They use monthly, physical, sales and inventory data collected by a private industry-specific business intelligence firm. New sources of such smaller scale data have the potential to greatly increase the possibilities of what is possible in empirical research.

The technological advances of the past two decades made it possible for firms to collect and store large amounts of information. Researchers, however, can only access this kind of data through individual collaborations with firms—limiting its adoption[3]. Ideally, large, high frequency databases of primary data will one day be able to replace the sources of aggregate information we have today—if firms are willing to share such information and support the publication of the insights that originate from its use. Such primary data can be of a material nature, thus allowing researchers to avoid the use of financial proxies, which (as shown by Chen et al., 2014) can lead to under/over estimations of performance metrics.

What is the dynamic effect of arbitrary changes to inventory buffers?

In large sections of this dissertation, we hypothesize that decision-makers have economic and financial incentives to alter their inventory buffers. However, no research has been done to study this analytically. in part because, from a theoretical perspective, parameter changes are not interesting when looking at steady state performance. However, as we have shown through this dissertation, the dynamic responses brought about by such changes are of extreme importance. Robustness in the dynamic response should be analyzed when adopting or altering policies. Of course, the study of the dynamic consequences of parameter changes is not without challenges. For example, extending control theoretic models to allow arbitrary changes in parameters would make them non-linear, which would preclude us from using the linear control theory frameworks developed during the past 30 years.

[3] Case in point, the research presented in Chapter 3 uses such primary data for the upstream echelons, but suffers the disadvantage of the non-availability of similar supply-chain-wide data.

Likewise, capacity limitations introduce a similar non-linearity into the models. Relaxing the assumption of an inventory-production system of infinite capacity is a hard problem, but nonetheless a practically relevant one.

What is the influence of capacity limitations on the bullwhip?

Capacity limitations, both at a production level (i.e. how much can I produce at one?) and at an inventory level (i.e. how much can I store at once?) are a major issue in real-life production systems. The inventory models presented in this dissertation assume that no capacity constraints exist in a given firm. This means that a firm will always fulfill orders within the lead time, and that a firm will always be able to store any arbitrary amount of product.

The reason why capacity limitations are potentially relevant to the study of the bullwhip is that, by definition, the bullwhip generates oscillations in both orders and inventories, whereas capacity limits impose a hard boundary to those oscillations. At first sight, this appears to imply that capacity limitations will keep the bullwhip in check[4], however, the theory suggests that in the presence of shortages, humans will *increase* orders in subsequent periods. This phenomena, called shortage gaming, appears in the hope that a larger order is reciprocated with a larger allocation of the available capacity. These allocation decisions, however, will depend on a multitude of factors—order size may or may not be one of them. Unfilled orders, nevertheless, are usually tracked by suppliers; shortage gaming, when adopted as a strategy to secure a large allocation, is usually followed by the *return of the bullwhip*: a surge in deliveries from suppliers catching up to unfilled demand (De Kok, 2012).

Thus, the effect of capacity limitations in a supply chain goes beyond an arbitrary bound on a model; it involves human behavior (shortage gaming) as well as firm policy (allocation rules). While expanding simulation models to include capacity limitations is trivial, this tells only part of the story. Further research into the influence of capacity limitations in real life can be performed using empirical data and the natural experiment of the 2008 crisis, for example.

Moreover, on the flip-side of capacity limitations issues experienced by suppliers, we find financing limitations experienced by customers. Another common assumption in inventory models is that customers have the necessary liquidity to finance their orders, no matter their size. In this dissertation, however, we have seen that liquidity is an important factor in the day to day running of firms.

[4]The reasoning being that if a bullwhip-related spike in my orders exceeds my supplier's available capacity, then my *actual* orders will be equal to said capacity and consequently, smaller than the original spike.

From a simplified point of view, customers' financial limitations can be seen to be akin to suppliers' capacity limitations: They bound the possible orders, but also introduce behavioral and firm dynamics into the picture. Unlike capacity, however, the financial standing of publicly traded companies is available for researchers to use.

Are there other possible sources of bullwhip that we can study?

The 2008 credit crisis, with its synchronization and magnitude, served as a powerful natural experiment. We expect that other similarly synchronized phenomena can be used for the study of the bullwhip—and for the application of taming measures. A recent example of such a natural experiment is the massive flood that struck Thailand in 2011. Production in large parts of the country was essentially halted.

From an empirical point of view, to study the way in which different firms (or sectors) respond to different "bullwhip events" can generate interesting insights. Particularly, we can test whether the rogue seasonality defined by the particular parameters of the system is observable in the successive responses. If this is true, then the structural footprint of different supply chains can be calculated and used to aid in the forecasting of future reactions.

In conclusion, it is clear that the Bullwhip Effect, and inventory dynamics in general, are a ripe field for future research whether it be analytical, empirical or experimental. The hypothetical research questions presented in this section are direct extensions of some of the limitations and simplifying assumptions present in this dissertation. Naturally, they do not represent an all-inclusive list—new research directions are limited by one's own imagination. They do, however, represent a series of concrete directions that are realistic and with potentially significant impact in the practical application of our knowledge.

One can argue that the reason we use mathematical relationships to model, and ultimately achieve a deeper understanding of the matter we study is, as physicist Richard Feynman put it, because *"[the real reason is that] the subject is enjoyable"*. However, when choosing research directions in the presence of a virtually infinite number of choices, we mustn't forget that, in our field, the potential for application of our results is of significance. Choose accordingly: fun and impactful research.

Bibliography

A. B. Abel, B. S. Bernanke, and D. Croushore. *Macroeconomics (6th edn)*. Boston: Pearson Education, Inc, 2008.

M. Abramovich and I. Stegun. *Handbook of mathematical functions with formulas graphs and mathematical tables.* 1965.

G. Alessandria, J. P. Kaboski, and V. Midrigan. The great trade collapse of 2008-09: An inventory adjustment? NBER Working Papers 16059, National Bureau of Economic Research, Inc, 2010.

H. I. Ansoff and D. P. Slevin. An appreciation of industrial dynamics. *Management Science*, 14(7):383–397, 1968. doi: 10.1287/mnsc.14.7.383.

R. Baldwin. *The great trade collapse: Causes, Consequences and Prospects*. Centre for Economic Policy Research, 2009.

R. M. Belchamber, D. Brister, M. P. Collins, T. J. Hill, R. N. R. Watkins, and D. G. Wood. Method for monitoring acoustic emissions., 1992.

D. A. Belsley. *Industry production behavior: The order-stock distinction*, volume 62. North-Holland, 1969.

W. J. Bertrand. Analysis of a production-inventory control system for a diffusion department. *International Journal of Systems Science*, 11(5):589–606, 1980.

A. S. Blinder. Inventories and sticky prices: More on the microfoundations of macroeconomics, 1983.

A. S. Blinder. Can the production smoothing model of inventory behavior be saved? *The Quarterly Journal of Economics*, January 1986.

A. S. Blinder. *Inventory theory and consumer behavior*. University of Michigan Press, 1990.

A. S. Blinder and L. Maccini. Taking stock: A critical assessment of recent research on inventories. *The journal of economic perspectives*, 5(1):73–96, 1991. ISSN 0895-3309.

E. H. Bowman. Consistency and optimality in managerial decision making. *Management Science*, 9(2):310–321, 1963.

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

R. L. Bray and H. Mendelson. Disentangling production smoothing from the bullwhip effect. *Available at SSRN*, 2013.

R. Brent. *Algorithms for minimization without derivatives*. Dover Publications, 2002.

X.-z. Bu, Z.-h. Zhou, and R.-f. Guo. An empirical study on bullwhip effect in the supply chain; based on the statistical data from manufacturing industry of china. *8th International Conference on Service Systems and Service Management (ICSSSM)*, pages 1–6, 2011.

J. F. Burns and B. Sivazlian. Dynamic analysis of multi-echelon supply systems. *Computers & Industrial Engineering*, 2(4):181–193, 1978.

G. P. Cachon and M. Fisher. Supply Chain Inventory Management and the Value of Shared Information. *Management Science*, 46(8):1032–1048, August 2000.

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

S. Cannella and E. Ciancimino. The APIOBPCS Deziel and Eilon parameter configuration in supply chain under progressive information sharing strategies. *Proceedings of the 40th Conference on Winter Simulation.*, pages 2682–2689, 2008.

C. Chatfield. The holt-winters forecasting procedure. *Applied Statistics*, 27(3):264, 1978.

F. Chen, Z. Drezner, and J. Ryan. Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management Science*, Vol. 46(No. 3):436–443, March 2000.

F. Chen. Information sharing and supply chain coordination. *Handbooks in operations research and management science*, 11:341–421, 2003.

L. Chen and H. L. Lee. Information sharing and order variability control under a generalized demand model. *Management Science*, 55(5):781–797, 2009.

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

L. Chen, W. Luo, and K. Shang. On the implications of measuring bullwhip effect with material flow data. *Working Paper*, 2014.

Y. F. Chen and S. M. Disney. The myopic Order-Up-To policy with a proportional feedback controller. *International Journal of Production Research*, 45(2):351–368, January 2007.

G. L. Childs and J. Johnston. *Unfilled orders and inventories: A structural analysis*. North-Holland, 1967.

A. J. Clark and H. Scarf. Optimal policies for a multi-echelon inventory problem. *Management science*, 6(4):475–490, 1960.

L. Cohen and A. Frazzini. Economic links and predictable returns. *The Journal of Finance*, 63(4):1977–2011, 2008. ISSN 1540-6261. doi: 10.1111/j.1540-6261.2008.01379.x.

R. Croson and K. Donohue. Upstream versus downstream information and its impact on the bullwhip effect. *System Dynamics Review*, 21(3):249–260, 2005. ISSN 1099-1727.

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

R. Croson, K. Donohue, E. Katok, and J. Sterman. Order stability in supply chains: Coordination risk and the role of coordination stock. *Production and Operations Management*, 2014.

A. G. De Kok. The return of the bullwhip. *Review of Business and Economic Literature*, 57(3), 2012.

A. G. De Kok and J. C. Fransoo. Planning supply chain operations: definition and comparison of planning concepts. *Handbooks in operations research and management science*, 11:597–675, 2003.

A. G. De Kok, F. Janssen, J. van Doremalen, E. van Wachem, M. Clerkx, and W. Peeters. Philips Electronics Synchronizes Its Supply Chain to End the Bullwhip Effect. *Interfaces*, 35(1):37–48, February 2005.

J. Dejonckheere, S. M. Disney, M. R. Lambrecht, and D. R. Towill. Transfer function analysis of forecasting induced bullwhip in supply chains. *International Journal of Production Economics*, 78(2):133–144, July 2002.

J. Dejonckheere, S. M. Disney, M. Lambrecht, and D. Towill. Measuring and avoiding the bullwhip effect: A control theoretic approach. *European Journal of Operational Research*, 147(3):567–590, 2003.

J. Dejonckheere, S. M. Disney, M. R. Lambrecht, and D. R. Towill. The impact of information enrichment on the bullwhip effect in supply chains: A control engineering perspective. *European Journal of Operational Research*, 153(3):727–750, March 2004.

M. Dekker, K. van Donselaar, and P. Ouwehand. How to use aggregation and combined forecasting to improve seasonal demand forecasts. *International Journal of Production Economics*, 90(2):151 – 167, 2004. ISSN 0925-5273. doi: http://dx.doi.org/10.1016/j.ijpe.2004.02.004.

D. Deziel and S. Eilon. A linear production-inventory control rule. *Production Engineer*, 46(2):93, february 1967.

E. Diehl and J. D. Sterman. Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes*, 62(2):198–215, 1995.

E. B. Diks and A. G. De Kok. Optimal control of a divergent multi-echelon inventory system. *European Journal of Operational Research*, 111(1):75–97, 1998.

S. M. Disney and D. R. Towill. A methodology for benchmarking replenishment-induced bullwhip. *Supply Chain Management: An International Journal*, 11(2): 160–168, 2006.

S. M. Disney, I. Farasyn, M. Lambrecht, D. R. Towill, and W. van de Velde. Taming the bullwhip effect whilst watching customer service in a single supply chain echelon. *European Journal of Operational Research*, 173(1):151–172, 2006a.

S. M. Disney. Supply chain aperiodicity, bullwhip and stability analysis with jury's inners. *IMA Journal of Management Mathematics*, 19(2):101–116, 2008. doi: 10.1093/imaman/dpm033.

S. M. Disney and D. R. Towill. On the bullwhip and inventory variance produced by an ordering policy. *Omega*, 31(3):157–167, 2003.

S. M. Disney, D. R. Towill, and R. D. H. Warburton. On the equivalence of control theoretic, differential, and difference equation approaches to modeling supply chains. *International Journal of Production Economics*, 101(1):194–208, May 2006b.

S. M. Disney, L. Hoshiko, L. Polley, and C. Weigel. Removing bullwhip from lexmark's toner operations. *POMS Meeting proceeding*, 2013.

K. Dooley, T. Yan, S. Mohan, and M. Gopalakrishnan. Inventory management and the bullwhip effect during the 2007-2009 recession: Evidence from the manufacturing sector. *Journal of Supply Chain Management*, 46(1):12–18, 2010. ISSN 1745-493X.

G. Dougherty. *Pattern Recognition and Classification: An Introduction.* Springer Science + Business Media, New York, 2013.

J. Edghill and D. R. Towill. Assessing manufacturing system performance: frequency response revisited. *Engineering Costs and Production Economics*, 19 (1):319–326, 1990.

M. S. Eichenbaum. Some empirical evidence on the production level and production cost smoothing models of inventory investment. *National Bureau of Economic Research*, Working Paper No. 2523, 1990.

H. Escaith, N. Lindenberg, and S. Miroudot. International supply chains and trade elasticity in times of global crisis. *World trade organization: Economic research and statistics division*, (Staff working paper ERSD-2010-08), 2010.

C. E. Fee, C. J. Hadlock, and S. Thomas. Corporate equity ownership and the governance of product market relationships. *The Journal of Finance*, 2006.

C. E. Fee and S. Thomas. Sources of gains in horizontal mergers: evidence from customer, supplier, and rival firms. *Journal of Financial Economics*, 74(3):423–460, 2004.

M. Feldstein and A. Auerbach. Inventory behavior in durable-goods manufacturing: The target-adjustment model. *Brookings Papers on Economic Activity*, 7(2):351–408, 1976.

T. Fitzgerald. Inventories and the business cycle: an overview. *Economic Review-Federal Reserve Bank of Cleveland*, 33:11–22, 1997.

J. Forrester. Industrial dynamics: a major breakthrough for decision makers. *Harvard Business Review*, 36(4):37–66, 1958.

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

P. Gao and H. Yun. Commercial paper, lines of credit, and the real effects of the financial crisis of 2008: Firm-level evidence from the manufacturing industry. *Notre Dame University*, 2009.

P. Garnell and D. J. East. *Guided weapon control systems*, volume 2. Pergamon Press Oxford, England, UK, 1977.

V. Gaur, A. Giloni, and S. Seshadri. Information sharing in a supply chain under arma demand. *Management Science*, 51(6):961–969, 2005. doi: 10.1287/mnsc.1050.0385.

S. Geary, S. M. Disney, and D. R. Towill. On bullwhip in supply chains—historical review, present practice and expected future impact. *Intern. Journal of Production Economics*, 101(1):2–18, 2006.

M. A. Ghali. Seasonality, aggregation and the testing of the production smoothing hypothesis. *The American Economic Review*, pages 464–469, 1987.

R. M. Goodwin. Secular and cyclical aspects of the multiplier and the accelerator. *Income, Employment and Public Policy: Essays in Honor of AH Hansen (various authors), New York: Norton*, 108:132, 1948.

M. F. Gorman and J. I. Brannon. Seasonality and the production-smoothing model. *International Journal of Production Economics*, 65(2):173 – 178, 2000. ISSN 0925-5273.

S. C. Graves. A single-item inventory model for a nonstationary demand process. *Manufacturing & Service Operations Management*, 1(1):50–61, 1999. doi: 10.1287/msom.1.1.50.

J. Hammond. Barilla spa. harvard business school case 9-694-046, 1994.

G. A. Hay. Adjustment costs and the flexible accelerator. *The Quarterly Journal of Economics*, 1970a.

G. A. Hay. Production, price, and inventory theory. *The American Economic Review*, 1970b.

K. Hoberg, J. Bradley, and U. Thonemann. Analyzing the effect of the inventory policy on order and inventory variability with linear control theory. *European Journal of Operational Research*, 176(3):1620–1642, 2007a.

K. Hoberg, U. Thonemann, and J. Bradley. Analyzing the effect of inventory policies on the nonstationary performance with transfer functions. *IIE Transactions*, pages 1–15, Jul 2007b.

C. Holt, F. Modigliani, J. Muth, and H. Simon. *Planning Production, Inventory and Work Force.* Prentice Hall, Englewood Cliffs, NJ., 1960.

J. B. Homer. Partial-model testing as a validation tool for system dynamics (1983). *System Dynamics Review*, 28(3):281–294, August 2012.

T. Hosoda and S. M. Disney. On variance amplification in a three-echelon supply chain with minimum mean square error forecasting. *Omega*, 34(4):344–358,

2006.

M. Jakšič and B. Rusjan. The effect of replenishment policies on the bullwhip effect: A transfer function approach. *European Journal of Operational Research*, 184(3):946–961, February 2008.

S. John, M. Naim, and D. Towill. Dynamic analysis of a wip compensated decision support system. *International Journal of Management Systems Design*, 1 4(4):283–297., 1994.

I. Jolliffe. *Principal Component Analysis*. John Wiley & Sons, Ltd, 2005. ISBN 9780470013199. doi: 10.1002/0470013192.bsa501. URL `http://dx.doi.org/10.1002/0470013192.bsa501`.

E. Jury. A note on aperiodicity condition for linear discrete systems. *Automatic Control, IEEE Transactions on*, 30(11):1100–1101, 1985.

E. Jury. *Theory and Application of the z-Transform Method*, volume 3. Wiley New York., 1964.

J. A. Kahn. Inventories and the volatility of production. 1986.

J. A. Kahn. 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, 1992.

D. Kahneman, A. Tversky, and P. Slovic. *Judgment under uncertainty: Heuristics and biases*. Cambridge university press, 1982.

T. Kelepouris, P. Miliotis, and K. Pramatari. The impact of replenishment parameters and information sharing on the bullwhip effect: A computational study. *Computers & Operations Research*, 35(11):3657–3670, November 2008.

I. Kim and M. Springer. Measuring endogenous supply chain volatility: Beyond the bullwhip effect. *European Journal of Operational Research*, 189(1):172–193, 2008.

R. Lai. Bullwhip in a spanish shop. *Harvard NOM Working Paper*, 06(06), 2005.

R. Lai. Is inventory's fiscal year end effect caused by sales timing? a test using a natural experiment from germany. Technical report, Working Paper, Harvard University, 2008.

H. L. Lee, V. Padmanabhan, and S. Whang. The bullwhip effect in supply chains. *Sloan management review*, 38(3):93–102, 1997a.

H. L. Lee, V. Padmanabhan, and S. Whang. Information distortion in a supply chain: the bullwhip effect. *Management science*, 43(4):546–558, 1997b. ISSN 0025-1909.

H. L. Lee, K. C. So, and C. S. Tang. The Value of Information Sharing in a Two-Level Supply Chain. *Management Science*, 46(5):626–643, May 2000.

A. A. Levchenko, L. T. Lewis, and L. L. Tesar. The Collapse of International Trade during the 2008–09 Crisis: In Search of the Smoking Gun. *IMF Economic Review*, 58(2):214–253, December 2010.

V. I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. In *Soviet physics doklady*, volume 10, page 707, 1966.

Q. Li, S. M. Disney, and G. Gaalman. Avoiding the bullwhip effect using Damped Trend forecasting and the Order-Up-To replenishment policy. *International Journal of Production ...*, 2013.

M. Lovell. Manufacturers' inventories, sales expectations, and the acceleration principle. *Econometrica: Journal of the Econometric Society*, 1961.

M. Lovell. Inventories, Production Smoothing, and the Flexible Accelerator. *The Quarterly Journal of Economics*, January 1971.

M. Lovell. Researching inventories: Why haven't we learned more? *International Journal of Production Economics*, 35:33–41, Jan 1994.

E. Lundberg. *Studies in the theory of economic expansion*. PS King & Son, ltd., 1937.

S. Makridakis and M. Hibon. The m3-competition: results, conclusions and implications. *International Journal of Forecasting*, 16(4):451 – 476, 2000. ISSN 0169-2070.

M. Marden. *The Geometry of the Zeros of a Polynomial in a Complex Variable*. American Mathematical Society, 1969.

R. B. Marr and G. H. Vineyard. Five-diagonal toeplitz determinants an their relation to chebyshev polynomials. *SIAM journal on matrix analysis and applications*, 9(4):579–586, 1988.

R. Metters. Quantifying the bullwhip effect in supply chains. *Journal of Operations Management*, 15(2):89–100, 1997. ISSN 0272-6963.

L. Metzler. The nature and stability of inventory cycles. *The Review of Economic Statistics*, 23(3):113–129, 1941. ISSN 0034-6535.

N. Nise. *Control Systems Engineering, (With Cd)*. John Wiley & Sons, 2007.

R. Oliva. Model calibration as a testing strategy for system dynamics models.

European Journal of Operational Research, 151(3):552–568, 2003.

R. Oliva and J. Sterman. Cutting corners and working overtime: Quality erosion in the service industry. *Management Science*, 47(7):894–914, Jan 2001.

M. Olivares, C. Terwiesch, and L. Cassorla. Structural estimation of the newsvendor model: An application to reserving operating room time. *Management Science*, 54(1):41–55, 2008. doi: 10.1287/mnsc.1070.0756. URL http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1070.0756.

M. Ortega and L. Lin. Control theory applications to the production–inventory problem: a review. *International Journal of Production Research*, 42(11):2303–2322, 2004.

M. Pesch and K. Hoberg. Inventory management under financial distress: an empirical analysis. *Working Paper: Kühne Logistics University*, 2013.

S. Raghunathan. Information Sharing in a Supply Chain: A Note on its Value when Demand Is Nonstationary. *Management Science*, 47(4):605–610, April 2001.

V. A. Ramey and D. J. Vine. Declining volatility in the US automobile industry. *The American Economic Review*, 2006.

V. A. Ramey and K. D. West. Inventories. *Handbook of macroeconomics*, 1:863–923, 1999.

D. F. Rogers and S. Tsubakitani. Newsboy-style results for multi-echelon inventory problems: backorders optimization with intermediate delays. *Journal of the Operational Research Society*, 1991.

S. Rumyantsev and S. Netessine. What Can Be Learned from Classical Inventory Models? A Cross-Industry Exploratory Investigation. *Manufacturing & Service Operations Management*, 9(4):409–429, January 2007.

C.-w. Shen. Temporal aggregation and the production smoothing model: Evidence from electronic parts and components manufacturing in taiwan. In A. Chan and S.-I. Ao, editors, *Advances in Industrial Engineering and Operations Research*, volume 5 of *Lecture Notes in Electrical Engineering*, pages 83–94. Springer US, 2008. ISBN 978-0-387-74903-7.

J. Shlens. A tutorial on principal component analysis. *Systems Neurobiology Laboratory, University of California at San Diego*, 2005.

V. Shukla, M. M. Naim, and N. F. Thornhill. Rogue seasonality detection in supply chains. *International Journal of Production Economics*, 138(2):254–272,

2012.

J. R. Silvester. Determinants of block matrices. *The Mathematical Gazette*, 84(501): 460–467, Nov 2000.

H. Simon. On the application of servomechanism theory in the study of production control. *Econometrica*, 20(2):247–268, Apr 1952.

L. Sprague and J. Wacker. Macroeconomic analyses of inventories: learning from practice. *International journal of production economics*, 45(1-3):231–237, 1996.

M. Springer and I. Kim. European Journal of Operational Research. *European Journal of Operational Research*, 203(2):380–392, June 2010.

S. Steinker and K. Hoberg. Does it swing? an empirical analysis of firm-level inventory volatility. *Working Paper: Kühne Logistics University*, 2012.

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

J. D. Sterman. *Business dynamics: Systems thinking and modeling for a complex world*. Irwin/McGraw-Hill, 2000. ISBN 007238915X.

J. D. Sterman. Does formal system dynamics training improve people's understanding of accumulation? *System Dynamics Review*, 26(4):316–334, 2010.

D. Struik. *A source book in mathematics, 1200-1800*. Harvard University Press, 1969.

S. Sudarsanam and J. Lai. Corporate financial distress and turnaround strategies: an empirical analysis. *British Journal of Management*, 12(3):183–199, 2001.

H. Theil. *Applied economic forecasting,*. North-Holland Pub. Co., Amsterdam, 1966.

N. F. Thornhill, S. L. Shah, B. Huang, and A. Vishnubhotla. Spectral principal component analysis of dynamic process data. *Control Engineering Practice*, 10 (8):833–846, 2002.

N. F. Thornhill and M. M. Naim. An exploratory study to identify rogue seasonality in a steel company's supply network using spectral principal component analysis. *European Journal of Operational Research*, 172(1):146–162, July 2006.

D. R. Towill. Dynamic analysis of an inventory and order based production control system. *The international journal of production research.*, 20(6):671–687,

1982.

D. Towill, Dennis R, L. Zhou, and S. Disney. Reducing the bullwhip effect: Looking through the appropriate lens. *International Journal of Production Economics*, 108(1):444–453, 2007.

J. R. Trapero, F. P. Garc, N. Kourentzes, et al. Impact of demand nature on the bullwhip effect. bridging the gap between theoretical and empirical research. In *Proceedings of the Seventh International Conference on Management Science and Engineering Management*, pages 1127–1137. Springer, 2014.

A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185:1124–1131, 1974.

H. Vassian. JSTOR: Journal of the Operations Research Society of America, Vol. 3, No. 3 (Aug., 1955), pp. 272-282. *Journal of the operations research society of America*, 1955.

P. Wang and Y. Wen. Inventory accelerator in general equilibrium. *Federal Reserve Bank of St Louis Working Paper*, 2009.

Y. Wen. Understanding the inventory cycle. *Journal of Monetary Economics*, 52(8): 1533–1555, November 2005.

A. Wunsch. *Complex variables with applications*. Addison-Wesley, 1983.

Summary

Inventory dynamics and the bullwhip effect: Studies in supply chain performance

More often than not, discussions regarding the bullwhip effect begin with diapers. This is not the exception. Some years ago, while analyzing the order patterns of their pampers line of diapers, Procter and Gamble executives were surprised to see that orders placed by distributors had considerably more variation than the retail sales to the public. In fact, this increase in variability also held when comparing orders and sales at different stages of the supply chain: The further upstream a firm was, the higher the variability of their orders. This phenomenon, it turns out, is not exclusive to diaper sales; it affects essentially every supply chain in the world. Today, we call it the bullwhip effect.

In general, the causes of the bullwhip effect are broadly separated into operational and behavioral. As the name implies, operational causes refer to factors around the way that firms implement their ordering processes. They are the updating of demand forecasts (due to the necessity to hedge against demand uncertainty), order batching (due to the tendency to buy less often in pre-defined multiples), price fluctuations (and the resulting speculative motives for inventory holding), and shortage gaming (due to customers inflating orders to gain better allocations in case of shortages).

Behavioral causes of the bullwhip effect, on the other hand, refer to those causes that stem from the innate humanness of decision making. Most notably, our apparent inability to keep track of cumulative quantities. In a supply chain context this is observed in the way that decision makers tend to under-estimate the supply pipeline. In other words, when calculating replenishment orders, we fail to fully take into account the inventory that has already been ordered but not yet received. This leads to the telltale over-ordering and over-adjusting of

the bullwhip effect, and has been replicated over and over by researchers all over the world—usually by running experiments where students manage simplified supply chains, where the operational causes of the bullwhip are controlled for.

In this thesis we investigate the dynamics of inventories and how they are intrinsically linked to the bullwhip effect. In particular, we explicitly model different mechanisms of the aforementioned under-estimation of the supply pipeline to study how this behavior, the bullwhip effect, and the dynamic nature of inventories are interrelated.

To understand exactly how prevalent and encompassing a subject the bullwhip effect is, we perform a survey that spans over 50 years of research, several disciplines, and multiple research methodologies. We trace the development of dynamic supply chain models to the 1950's and show how they led to the establishment of an new discipline –system dynamics– capable of simulating supply chain dynamics using the limited computing power of the age. In fact, we find that Forrester, already in 1958, describes the bullwhip effect in terms of business cycles. Using system dynamics, he describes the mechanisms that lead to the amplification of orders variance, as well as the appearance of a "fake business cycle": The emergence of a seasonality in orders that is not contained in the demand, an effect which today we know as rogue seasonality.

Among other research directions, we encounter, relevant to this thesis, the aforementioned separation of the causes of the bullwhip effect into operational and behavioral parameters and, perhaps surprisingly, substantial debate regarding its measurement and appearance on empirical data at different aggregation levels. For all the body of work available, the role of inventories is often implied and taken for granted: The view of economists, often assuming them to be an adjustment variable, clashing with the OM/OR view of inventories as explicit drivers of ordering decisions. In this view, inventories are the medium through which the bullwhip propagates, but their direct influence on the bullwhip dynamics are seldom analyzed.

To understand the impact of inventory decisions on the supply chain dynamics, we develop a system dynamics model based upon behavioral operations methodologies and use it to model four real-life supply chains at an echelon level of aggregation. We use the recent 2008 financial crisis as natural experiment to validate our models with primary data collected at a dutch chemical company.

Our study shows that the use of inventories as an explicit financial decision-making instrument (de-stocking to turn inventories into cash and increase liquidity) can act as a trigger to an inventory-driven bullwhip effect that accurately predicts the wave-like demand dynamics experienced by manufacturers during this period of economic distress.

In addition to this contribution to theory, we add to practice by demonstrating that these insights can be used operationally for demand predictions. We develop a prediction method by which final consumer demand, potentially four to five levels down the supply chain, is taken as the only exogenous information, whereupon the system dynamics models are used to propagate the demand upstream.

While the exogenous demand at consumer level drives the overall demand evolution, short-term demand dynamics are mainly driven by endogenous ordering decisions in the supply chain and the de-stocking response to the crisis. To increase the confidence in our de-stocking hypothesis, we test the models in the absence of the inventory shock induced by the synchronized de-stocking, and find that while they show good tracking of the long term evolution, they cannot replicate the magnitude and timing of the observed upstream demand.

Furthermore, the empirical evidence presented shows that slow reaction speeds and under-estimation of the supply pipeline correctly explain firm and echelon level observations. At this level, however, the pipeline under-estimation is not necessarily an entirely behavioral phenomenon. It is potentially a cause of a combination of the inherent reaction times of firms, and a decision rule that primarily steers, based upon, the on-hand inventory level.

Motivated by these findings, we perform additional studies to further our understanding of the interaction between inventories, behavior, and the bullwhip.

Using control theory, we develop a single-echelon model that allows us to study the inventory and order dynamics in an analytical way, explicitly modeling the behavioral mechanisms. In terms of theory, we develop a closed form expression that allows us test for the stability of the system as a function of its behavior, as measured by the over- or under-estimation of the supply pipeline. A stable system, in a supply chain context, guarantees that the orders generated by the policy will be finite. Additionally, we characterize the trade-off between the dynamic and stationary performance of the system and find that performance does not solely depend on the behavior itself, but rather on the way that behavioral parameters interact with different demand patterns. Behavior that is advantageous in the presence of cyclical demands is detrimental in the presence of shocks, and vice versa. The results of this study suggest that our prior observations are aligned with a behavior that –instead of being optimal for any one type of demand– seeks to control the bullwhip in the presence of changing demand patterns.

To further understand this interaction between behavior and demand, we perform extensive numerical experiments based upon the control theoretic model. We investigate the effect of behavior on performance by measuring the system's rogue seasonality. We find that the behavioral mechanisms can drive the seasonality of the system –overpowering the original demand seasonality– and that this resulting pattern is different for orders and inventories. With this study we add a new dimension to the study of "bullwhip-optimal" policies: The amount of over- or under-estimation of the supply pipeline determines a natural frequency for the inventory and order responses which can, depending on the particular behavior, be substantially different. The consequence of these observations is immediately perceived when we study the interaction between the demand's seasonality and the system's natural frequency. When the demand has a high frequency seasonality, over-estimating the supply pipeline helps reduce the bullwhip of the system. When the demand is of a low frequency, the opposite is true.

From an operational perspective, this implies that characterizing and tracking the seasonality of customer demand is of importance inasmuch as it defines the core response of the system. In addition to understanding the seasonal component of demand, our study calls for an understanding of the natural response frequency of one's system. At a tactical level, this understanding allows managers to better understand the medium- to long-term evolution of inventories and orders, potentially affecting the way internal performance metrics work. In this view, the baseline from which to measure inventory performance is not stationary, but cyclical.

Finally, we focus on a different dimension of the relationship between inventories and the bullwhip: Are inventories a consequence, or are they drivers, of production decisions? Earlier in the thesis we tested the hypothesis that inventories triggered a bullwhip effect following the financial crisis. We did so by comparing the results of a model with and without inventory reductions. We now use historical financial data on approximately 6.000 distinct customer-supplier dyads to statistically test a series of hypothesis relating the upstream consequences of downstream inventory decisions, and the systematic components that influence such decisions. In the first part of this empirical study –covering 30 years of data– we find evidence of suppliers over-reacting to their customers' inventory changes. To understand the underlying reasons behind the inventory decisions, we impute a rational ordering behavior to upstream decision makers, and find that the observed data is consistent with the usage of inventories as a financial decision-making instrument, as hypothesized in the first of our studies. Given that the data analyzed contains the turbulent periods of the 2008 financial crisis, which could skew the results, we repeat our statistical analysis with sub-samples of the data. After separating the pre- and post-crisis data, we still find evidence of the overreaction of upstream orders to downstream inventory changes, and of systematic adjustments to inventory-related costs tied to the financial and economic conjuncture. Furthermore, when performing statistical tests on the effects, we find that the crisis period attenuates the former and exacerbates the latter.

About the author

Maximiliano Udenio was born in Buenos Aires, Argentina, on March 30th, 1982. He attended Belgrano Day School, and graduated with a high school diploma in December 1999. He received the title of Industrial Engineer from the Instituto Tecnológico de Buenos Aires (ITBA) in 2007. After a few years working in the photography business he moved out of the southern hemisphere and into Europe. Specifically to The Netherlands, where in June 2008 he started as a trainee in the Logistics Management Systems program of the Eindhoven University of Technology under the supervision of prof.dr.ir Jan. C. Fransoo.

The company project, performed as part of the LMS program, prompted the start of a PhD project to better understand and investigate what they had learned. From June 2010 until May 2014 he was a PhD Student at the TU/e, also under the supervision of professor Fransoo. Between September and December of 2013, Maximiliano visited Cornell University, in beautiful Ithaca, NY, to work with prof.dr Vishal Gaur on what would later become Chapter 6 of this thesis.

As of June 2014 he is working as an assistant professor at the Eindhoven University of Technology, in the Operations, Planning, Accounting and Control group. He defends his PhD thesis on September 25th, 2014.