8-2013

Data Science in Supply Chain Management: Data-Related Influences on Demand Planning

Yao Jin
University of Arkansas, Fayetteville

Data Science in Supply Chain Management: Data-Related Influences on Demand Planning

Data Science in Supply Chain Management: Data-Related Influences on Demand Planning

A dissertation submitted in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy in Supply Chain Management

By

Yao Jin
Centenary College of Louisiana
Bachelor of Science in Finance, 2005
University of Arkansas
Master of Business Administration, 2006

August 2013
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Dr. Matthew Waller
Dissertation Co-Director

Dr. Brent Williams
Dissertation Co-Director

Dr. Christian Hofer
Committee Member

ABSTRACT

Data-driven decisions have become an important aspect of supply chain management. Demand planners are tasked with analyzing volumes of data that are being collected at a torrential pace from myriad sources in order to translate them into actionable business intelligence. In particular, demand volatilities and planning are vital for effective and efficient decisions. Yet, the accuracy of these metrics is dependent on the proper specification and parameterization of models and measurements. Thus, demand planners need to step away from a "black box" approach to supply chain data science. Utilizing paired weekly point-of-sale (POS) and order data collected at retail distribution centers, this dissertation attempts to resolve three conflicts in supply chain data science. First, a hierarchical linear model is used to empirically investigate the conflicting observation of the magnitude and prevalence of demand distortion in supply chains. Results corroborate with the theoretical literature and find that data aggregation obscure the true underlying magnitude of demand distortion while seasonality dampens it. Second, a quasi-experiment in forecasting is performed to analyze the effect of temporal aggregation on forecast accuracy using two different sources of demand signals. Results suggest that while temporal aggregation can be used to mitigate demand distortion's harmful effect on forecast accuracy in lieu of shared downstream demand signal, its overall effect is governed by the autocorrelation factor of the forecast input. Lastly, a demand forecast competition is used to investigate the complex interaction among demand distortion, signal and characteristics on seasonal forecasting model selection as well as accuracy. The third essay finds that demand distortion and demand characteristics are important drivers for both signal and model selection. In particular, contrary to conventional wisdom, the multiplicative seasonal model is often outperformed by the additive model. Altogether, this dissertation advances both theory and practice in data science in supply chain management by peeking into the "black box" to identify several levers that managers may

control to improve demand planning. Having greater awareness over model and parameter specifications offers greater control over their influence on statistical outcomes and data-driven decisions.

ACKNOWLEDGMENTS

I could not have completed my studies and this dissertation without the love and support of many people around me. First and foremost, I would like to thank my mentors on the dissertation committee. I met Professor Matthew Waller in 2006 during my studies in the MBA program. My conversation with him planted the seeds which later sprouted into my main streams of research today. Together, Professor Waller and Professors Brent Williams and Christian Hofer provided me with invaluable guidance in my studies. In particular, I would like to truly thank them for their saintly patience as I learn how to write rigorous academic research, including this dissertation. I will always strive to reach the high standards they have consistently set for me throughout my studies. In addition, although Professor Adriana Hofer was not part of my dissertation committee, she frequently provided guidance and advice to me on both a professional and a personal level.

Secondly, I would like to sincerely thank all my loved ones. My girlfriend, Josie, offered me love and support. She is a major source upon which I drew energy and strength throughout my program. Moving from a bustling metropolis to a sleepy college town can be jarring, but she never wavered in her encouragement. I truly appreciate the vital role she played in supporting my spirits throughout my studies. My parents, Duan Li and Mingfeng Jin, came to the United States in 1993 and created a loving environment for me. As a first generation immigrant family, we were far from being wealthy. My parents scraped and saved every penny in order to create an environment conducive to academic excellence. Throughout all our hardship, their love and guidance made sure that I never deviated from a path of integrity and excellence.

Thirdly, I would like to especially thank my alma mater, Centenary College of Louisiana, for taking a chance in 2001 on a high school graduate who accomplished nothing particularly extraordinary. Although my high school record was above average and without many extra-curricular and leadership activities, Centenary College chose to bestow upon me, after multiple rounds of interviews, their highest academic scholarship, over many other applicants with far more illustrious high school records. The liberal arts education I received nurtured my intellectual curiosity and critical thinking skills that continue to have significant impact on me even today. In particular, Professors Helen Sikes and Barbara Davis gave me an early glimpse into a prospective career in academia through working with me to co-author multiple research papers for presentation at national conferences. I am forever indebted to them for their early mentorship.

I would like to also thank Mr. Wes Kemp and his wife, Mrs. Sharon Kemp. Their generosity and philanthropic support to the college and the department played a vital role in allowing me to concentrate on my studies and to explore the academic horizons. Very few people possess the depth of expertise in their chosen field and yet remain so humble. My conversations with Wes have had a profound impact on me. I will always strive to instill the same qualities of diligence, integrity, generosity, and tolerance, into my students.

Last but not least, many other individuals have contributed to my growth as a person and as an academic. Professor Stan Fawcett and his wife, Amydee Fawcett, frequently encouraged me and offered me honest and heart-to-heart advice. They also make awesome brownies. In addition, I would like to thank all my mentors from my days working for Walgreens: Mr. Wayne Box, Mr. Scott Dilley, and Mr. Brad Ulrich. Finally, I would like to give credit to my friends from Walgreens, Wesley Dillard, John Kelley, Jeff Sharum, and Guadalupe Chavez, for having many discussions with me over how to improve inventory management and demand planning processes, which continue to drive my research interest to this day.

DEDICATION

This dissertation is dedicated to my mother and father. We moved to the United States when I was ten years old. Throughout my years in middle school and high school, we never had much material wealth. However, I never felt like I was unloved. My mother and father tried their best to provide a loving and supportive environment for me. I remember our first meal at a fast food restaurant—and it was considered a treat. They saved money whenever possible to make sure that I can have at least most of the expected components of a "normal" childhood: toys, books, and even video games. I remember a particularly string of devastating financial hits that left us without a car, which is a rather unthinkable predicament in rural Texas. Each day when arrived home after my two-mile walk under the blistering Sun, my parents always made sure we had ice cream stocked in the freezer—another treat for a struggling first generation immigrant family.

There were times when I acted as expected of a typical rebellious teenager, but my mother and father were always willing to overlook the little things and not micromanage me. They were both in school full-time pursuing their respective second master's degree. Their responsibilities as teaching and research assistants kept them constantly preoccupied and afforded them little time to supervise me. Instead, they told me to always be salient of the consequences of my actions and to always hold myself accountable. Although they did not provide much input or guidance to me in my eventual field of study, my parents' decision to move to the United States in pursuit of a better life was the genesis of what I am able to accomplish today. No matter the hardship, their unconditional love, support, and guidance kept me away from the dangerous allures of alternate paths in life chosen by many peers around me. That is why this dissertation is dedicated to my parents and their sacrifices for me throughout the years, and to a broader extent, our pursuit of happiness and the American Dream.

TABLE OF CONTENTS

Chapter 1 – INTRODUCTION

I. INTRODUCTION

A. Background

Since the creation of decision support systems in the 1960s, data-driven decisions have become a key capability for obtaining and sustaining competitive advantage. Today, every transaction at many levels is collected in some form and stored in a database for decision-making purposes. By 2021, firms are expected to have accumulated over 35 zettabytes of data generated from activities throughout the supply chain (Cognizant, 2012). Processing and storing such high volumes of data can command high levels of resources within an organization. Wal-Mart, for example, collects detailed data on every single transaction receipt for every single customer. Their data can be detailed down to the exact stock-keeping unit (SKU), its quantity purchased and price, store location, register, and time. Over the course of a day, Wal-Mart collects as many as 24 million transactions to be stored in its database of 2.5+ petabytes (McCarthy, 2012).

Streams of literature in various fields have proposed that the process of collecting and analyzing business data to formulate and disseminate actionable intelligence is vital to firm competitive advantage. In marketing, for example, successful firms tend to be more adept at generating, disseminating, and responding to market intelligence (Kohli & Jaworski, 1990). In a supply chain, data transmitted through interorganizational information systems such as electronic data interchange (EDI) is the language through which firms communicate and coordinate joint actions (Hill & Scudder, 2002). Reliance on such firm strategies has only increased with technological advances, such as radio-frequency identification (RFID) and geo-cache data generated through global positioning systems (GPS) embedded in consumer electronic devices. Anecdotal exemplars such as Amazon.com’s successful customer segmentation efforts only give companies greater incentive to install even more sensors throughout the supply chain to amass data with ever-increasing greater levels of detail and volume.

With greater volumes of data collected at finer levels, resource intensity and automation required for data storage, processing and analysis also increases. As a result, companies need to clearly specify parameters for data input prior to automated analyses and joint decision-making. Moreover, firms facing rapid gains in data detail and volume remain largely without guidance with regard to the proper use of data. For example, suppliers may gain visibility to both customer and downstream demand signals through highly costly investments. These data can be analyzed on many levels of aggregation using a diverse set of quantitative models. Clearly, understanding the complex relationship among demand signals, data aggregation, and seasonal forecasting models is an important factor in maximizing the value of both the capital and relational investment made to enable information sharing.

In the retail supply chain, many firms engage in strategies such as sales and operations planning (S&OP) and collaborative planning, forecasting, and replenishment (CPFR) through information sharing (Yao & Dresner, 2008). For example, Wal-Mart leverages its transaction data to formulate myriad decisions ranging from predicting consumer sentiments to arranging both internal distribution and coordinated replenishment with external suppliers (Bollier, 2010). Yet, while synchronized decisions can improve integration (Olivia & Watson, 2011) and operational performance (Barratt & Barratt, 2011), supply chain partners often encounter difficulties in demand planning. Facing multiple sources of demand signals, demand planners are often at a loss in selecting the appropriate source and format of the demand signal used to forecast customer demand.

First, different functions within each firm operate under varying levels of aggregation (Pauwels et al., 2004). As a result, dominant functions within each firm tend to set the level of aggregation for analyses such as forecasting (McCarthy et al., 2004). Misguided attempts at remediating conflicting levels of aggregation, such as decentralized demand planning systems, can result in more harm than good (McCarthy et al., 2004). Those firms that adopt one-number forecasting would either utilize disaggregated data to make decisions at the aggregate level (i.e., bottom-up approach) or utilize aggregate data to make decisions at the disaggregate level (i.e., top-down approach). The degree of complexity is exacerbated when functions from different firms attempt to collaborate and share information taken from different levels and time buckets. The impact of misaligned data aggregation and levels of decision can result in inaccurate measurements and suboptimal decisions (Zotteri & Kalchschmidt, 2007), thereby compromising information relevancy as well as the effectiveness of resource utilization.

Second, the need to automate data processing rises along with the volume of data. Specific to forecasting, myriad quantitative models exist for seasonal and non-seasonal data (Makridakis et al., 1982; Makridakis & Hibon, 2000). Whereas it is fairly simple for firms to identify a priori a data series is seasonal or non-seasonal (Chatfield and Yar, 1988), the decision to use the proper seasonal forecasting model is much more ambiguous. Furthermore, with increased adoption of POS-sharing in the retail supply chain, suppliers have to consider not only the proper seasonal forecasting model but also whether or not to use POS to forecast customer orders. Considering the roles of bullwhip and the mathematical differences in the additive and multiplicative seasonal factors, choosing the wrong combination of information source and seasonal forecasting model can inflate forecast error to lead to demand planning conundrums.

To obtain a greater understanding of the effective use of supply chain data to generate actionable business intelligence, the goal of this dissertation is to diagnose the effect of data aggregation in supply chain management to facilitate greater accuracy in measuring and forecasting supply chain outcomes. Furthermore, this dissertation also attempts to gain additional insight to the effect of demand signal distortion on the accuracy and model selection for seasonal customer demand forecast.

B. Theoretical Background

Information sharing is an important tool for supply chain integration toward improved performance (Christopher 1997; Frohlich & Westbrook 2001; Allred et al. 2011). Considering that supply chain partners expend significant time and resources to establish both formal and informal linkages to facilitate collaborative efforts, many firms struggle to reap fruits of their investment (Jin et al. 2013). Even as firms continuously invest in information-heavy strategies (Ravichandran & Liu, 2011), substantial disconnect remains between collecting and utilizing information collected and shared by supply chain partners. This may be partly attributed to differences in the way firms aggregate their data due to a combination of mistaken beliefs as well as functional and practical constraints.

Data can be aggregated by product-location and by time (temporal). Under product-location aggregation, two distinct demand series defined either by product or by location are combined to form a single series. Under temporal aggregation, one demand series for one product is aggregated from a lower level of consecutive time units (e.g., weekly) to a higher level (e.g., monthly). Managerially, the motivation to aggregate such data can be either to reduce the overall amount of data for ease of use by a desktop workstation, or to match a level of analysis as needed by another business unit or firm. As a result, aggregation can benefit a firm by expediting data processing time for rapid decision-making and lessening the firm's IT hardware requirement for capital savings. On the other hand, aggregation can also obscure data's true underlying statistical process to present a skewed view on supply chain performance and customer demand.

Under both forms of aggregation, certain statistical properties are transformed (e.g., Amemiya & Wu 1972) to result in a statistical masking effect (e.g., Chen & Lee 2012). This effect occurs primarily due to different levels of stochastic variance in each disaggregated data series that offset when aggregated. Rossana & Seater (1995) study the effect of temporal aggregation and conclude that aggregation results in altered cyclical properties of subsets of time series data. In addition, Chen and Lee (2012) study the effect of both product-location and temporal aggregation on the measurement of the bullwhip effect and find that the aggregated view masks the true degree of demand distortion at disaggregated levels.

The consequences of statistical aggregation can result in conflicting conclusions. For example, the prevalence of the bullwhip effect is questioned in recent empirical literature (e.g., Cachon et al., 2007). Bray and Mendelson (2012) explain that product-location aggregation results in casting stochastic amplification and seasonal smoothing as two opposing forces simultaneously pulling the bullwhip ratio. Aggregation tends to result in greater emphasis on the seasonal variance, which can be easily smoothed to dampen the overall magnitude of the bullwhip effect.

In forecasting, Amemiya and Wu (1972) and Rossana and Seater (1995) analytically identified various transformative properties of the temporal aggregation process that results in altered and rogue statistical effects (i.e., information loss effect). They contend that forecast accuracy can be substantially compromised as a result. A subsequent body of analytical literature further expanded the list of statistical processes covered by the information loss effect. In contrast, Hotta et al. (2005) argue that temporal aggregation is instead beneficial to forecast accuracy due to the overall variance reduction effect, as extreme highs and lows become offset, to provide a time series that is less susceptible to outliers' effects.

Finally, while the use of downstream demand signal for forecasting is a widely-prescribed strategy for improving customer demand planning, Williams and Waller (2010) found that POS is beneficial only 65% of the time for retail order forecasts. Downstream demand signals benefit customer demand planning because it is free of the distortionary effect due to managerial and behavioral idiosyncrasies (Lee et al., 1997; Metters, 1997). On the other hand, while some idiosyncratic behaviors are unpredictable, retail inventory management policies associated with seasonal smoothing result in ordering patterns that possess cyclical variance that deviate from consumer behavior (Parkany, 1961). Therefore, different sources of demand signals also have consequences on the choice of two typical seasonal forecasting methods: Holt-Winter's additive and multiplicative models. Whereas the additive model assumes seasonality to be relatively constant, the multiplicative model assumes it to be proportional to mean demand (Chatfield & Yar, 1988).

Given the above theoretical conflicts, this dissertation attempts to reconcile the following: 1) What is the role of data aggregation in the conflicting empirical observance in the magnitude of the bullwhip effect? 2) Under what conditions is temporal aggregation beneficial or harmful to customer demand forecasts? 3) What are the drivers of seasonal forecasting accuracy and model selection? Furthermore, this dissertation will examine demand planning topics with a particular emphasis on two ubiquitous issues in supply chain management: Information-sharing and demand distortion.

C. Business Applications

The data from this study are collected from a large national consumer packaged-goods company. Data are collected for twenty-four products in three categories from ten DCs of a leading retailer. All three categories are frequently shopped and each in a different stage of category life cycle. In addition, each category has distinct shelf life (short, medium, and long), with one category being seasonal. Altogether, these three categories may be considered representative of most category demand characteristics for generalizability.

Each retailer DCs serves approximately one-hundred stores. As transactions occur at the store level, point-of-sale demand data are electronically transmitted to the DC for replenishment and operations purposes. Each week, DCs would generate orders based on the collective point-of-sale demand for all stores served. The orders are transmitted to the supplier for fulfillment. Delivery time generally had minimal impact on ordering policy. In addition, the retailer also shares with its supplier the point-of-sale data to assist them with capacity planning decisions.

A key research question is the effect of data aggregation. With greater data volume, data processing and analysis have become increasingly more difficult on the typical computer workstation. Thus, companies face the choice of either investing in greater information technology equipment to expand their capabilities, or to aggregate data and “shrink” the total size of the data down to a more manageable size. Moreover, functions and firms operating under conflicting time buckets and organizational hierarchies require different levels of analyses as well. Thus, this dissertation first examines the statistical effect of data aggregation on supply chain metrics.

Specifically, demand planning requires accurate assessment of demand uncertainty. Statistical effects of aggregation result in the offsetting of the highs and lows of data at the disaggregate level. As a result, the supplier's view on the true underlying demand volatility becomes obscured. Without an accurate view on demand uncertainty, it is likely that the supplier will be unable to make optimal decisions in anticipation of future demand and necessary capacity to fulfill customer requirements.

In addition, demand planning also relies on accurate demand forecast. Collaborative demand planning strategies such as sales and operations planning (S&OP) tend to follow one-number forecasting (Finn, 2004). That is, collaborating partners jointly decide on a single forecast for synchronized activities for efficient and effective replenishment and distribution. But not all companies operate on the same levels of aggregation. Whereas the supplier in this study replenishes its retail customer's DCs on a weekly basis, the retailer often makes decisions on a daily basis. Thus, this dissertation attempts to analyze and empirically test the impact of temporally aggregation on forecast accuracy.

Lastly, despite that POS-sharing is viewed as a vital method to improve forecasting, evidence indicates that suppliers should not always use POS as their forecast basis. In particular, seasonal products tend to have significant demand spikes during peak selling season. As a result, retail order policies for seasonal products emphasize operations smoothing rather than quickly responding to seasonal demand fluctuations. To further complicate seasonal demand forecasting, suppliers are largely without guidance as to which of the two commonly utilized seasonal forecasting models—Holt Winter's additive and multiplicative models—to use for each demand signal. Thus, demand signal and seasonal forecasting model selection is the third business problem this dissertation attempts to investigate.

D. Contributions

Christopher (1997) argued that competition is shifting from between companies to between supply chains. With increased information flow, supply chain partners are tasked with utilizing increasingly larger volume of data to obtain superior performance. Yet companies are also subject to resource and process constraints. Information technology is costly to acquire (Ravichandra & Liu, 2011) and difficult to deploy (Yao et al., 2012). Mismatched levels of temporal aggregation and methods for supply chain performance measurement and demand forecasting can potentially limit the effectiveness and returns on such costly projects.

Results from the first study suggest that data aggregation can obscure managers' view on bullwhip's prevalence and magnitude within the supply chain. Specifically, statistical effects of the aggregation process can mask the bullwhip effect at lower levels of aggregation. As a result, managers are left with a highly optimistic view of the underlying supply chain volatility. Considering the importance of accurately assessing bullwhip in demand planning, the level of measurement in terms of both product-location as well as temporal dimensions should be carefully considered to match the level of decision (Zotteri & Kalchschmidt, 2007). Moreover, the first study also demonstrates that substantial differences exist among bullwhip measures. While the ratios of coefficient of variation (Fransoo and Wouters, 2000) and variance (Lee et al., 1997) are relatively similar, the fractional growth rate (Cachon et al., 2007) shows significant potential to present a distorted view of bullwhip at disaggregate levels of analysis.

The second study investigates the effect of temporal aggregation on forecast accuracy. The study begins by exploring a body of literature that analyzes the statistical effects of temporal aggregation. The analytical stream argues that temporal aggregation is detrimental to forecasting due to information loss (e.g., Amemiya & Wu, 1972; Rossana & Seater, 1995). But a substantial stream of empirical literature concludes that temporal aggregation is actually beneficial to forecast accuracy due to variance reduction (e.g., Hotta et al., 2005). By conducting a quasi-experiment on forecast accuracy, study 2 finds that both statistical effects exist concurrently in temporal aggregation. Therefore, temporal aggregation can be selectively deployed as a tool to improve forecast accuracy. However, change to forecast accuracy due to temporal aggregation can be either positive or negative depending on the availability of downstream demand signals and the autocorrelation factor of the forecast input.

Study three utilizes seasonal POS and order data, to examine two issues in statistical forecast for seasonal demand. The first issue explored is the impact of demand signal selection on forecast error. Contrary to conventional belief, shared demand signals from downstream along the supply chain are not always superior to those readily observed by the supplier. A main reason for this is because seasonal products require substantial retailer intervention in order to smooth operations needs associated with seasonal demand spikes. Therefore, suppliers forecasting future customer (e.g., the retailer) demand should incorporate past retail ordering patterns as predictors of future replenishment needs. However, bullwhip once again distorts this finding. Due to heightened stochastic variance, retail orders suffering from increased levels of bullwhip can induce substantial forecast error, thereby making POS data a potentially superior source of demand signal. The second issue explored is associated with the various conditions that determine the appropriate seasonal forecast model to be utilized. Notably, results indicate that the likelihood of the multiplicative model outperforming additive increases with heightened bullwhip, the reverse is true if demand planners were to utilize POS as the forecast input.

Altogether, this dissertation investigates common practices in the specification and utilization of supply chain data to advance both theoretical and managerial understanding of best practices in generating actionable business intelligence. The dissertation finds that the data science in supply chain management for demand planning rely on recognizing demand distortion and characteristics, and selecting the optimal level of data aggregation.

E. Structure of the Dissertation

Moving forward, the rest of this dissertation will be structured as follows. First, a literature review will be presented to summarize three streams of research relevant to the dissertation: the theoretical and empirical literature on the bullwhip effect; the analytical and empirical literature on data aggregation; and finally, the modern demand forecasting environment in the retail supply chain. Immediately following literature review, study 1 will show that data aggregation has a direct impact on the measurement of bullwhip. We then move to study 2, in which the specific impact of temporal aggregation on forecast error is examined through a quasi-experiment in forecasting. Next, study 3 tests the impact of both demand signal and distortion on seasonal forecast accuracy and also conducts an exploratory analysis on their impact on seasonal forecasting model choice. Finally, the dissertation ends with an overarching conclusion as well as future opportunities for research.

F. REFERENCES

Allred, C.R., Fawcett, S.E., Wallin, C., & Magnan, G.M. (2012). A dynamic collaboration capability as a source of competitive advantage. *Decision Sciences*, 42, 129-161.

Amemiya, T., & Wu., R.Y. (1975), The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association*, 67 (39), 628-632.

Barratt, M., & Barratt, R. (2011). Exploring internal and external supply chain linkages: evidence from the field. *Journal of Operations Management,* 29, 514-528.

Bollier, D. (2010*) The Promise and Peril of Big Data*. The Aspen Institute, Washington DC.

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

Cachon, G. P., Randall, T., & Schmidt. G.M. (2007). In search of the bullwhip effect. *Manufacturing and Service Operations Management*, 9 (4), 457-479.

Chatfield, C. & Yar, M. (1988). Holt-Winters forecasting: some practical issues. *Journal of the Royal Statistical Society*, 37 (2). 129-140.

Chen, L., & Lee, H.L. (2012). Bullwhip effect measurement and its implications. *Operations Research,* 60, 771-784.

Christopher, M. (1997). *Marketing Logistics*, Butterworth-Heinemann, Oxford.

Cognizant, (2012). Big data's impact on the data supply chain. *20-20 Insights*, May.

Finn, R. (2004). Weekly vs monthly forecasting in the supply chain. *Logistics and Transport Focus,* 6, 22-28.

Fransoo, J.C., & Wouters, M.J.F. (2000). Measuring the bullwhip effect in the supply chain. *Supply Chain Management: An International Journal*, 5(2), 78-89.

Frohlich, M.T., & Westbrook, R. (2001). Arcs of integration: an international study of supply chain strategies. *Journal of Operations Management*, 19 (2), 185-200.

Hill, Craig A. & G.D. Scudder (2002) "The use of electronic data interchange for supply chain coordination in the food industry," *Journal of Operations Management*, (20), 375-387.

Hotta, L.K., Valls Pereira, P.L., & Otta, R. (2005). Effect of outliers on forecasting temporally aggregated flow variables. *Sociedad de Estadistica e Investigacion Operativa Test*, 13, 371-402.

Jin, Y., Fawcett, A.M., & Fawcett, S.E. (2013). Awareness is not enough: commitment and performance implications of supply chain integration. *International Journal of Physical Distribution & Logistics Management*, 43 (3), 205-230.

Kohli, A.K., & Jaworski, B.J. (1990). Market orientation: the construct, research propositions, and managerial implications. *Journal of Marketing*, 54 (2), 1-18.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43 (4), 546-559.

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: results of a forecasting competition. *Journal of Forecasting*, 1, 111-153.

Makridakis, S., & Hibon, M. (2000). The M3-competition: results, conclusions and implications. International *Journal of Forecasting*, 16, 451-476.

McCarthy, Erin (2012), The 5Ws and 1H of Big Data (Part 2 of 2). Retrived from http://www.sourcelink.com/blog/guest-author-series/2012/08/18/the-5ws-and-1h-of-big-data-(part-2-of-2)**.**

McCarthy, T.M., Davis, D.F., Golicic, S.L., & Mentzer, J.T. (2006). The evolution of sales forecasting management: a 20-year longitudinal study of forecasting practices. *Journal of Forecasting*, 25, 303-324.

Olivia, R., & Watson, N. (2011). Cross-functional alignment in supply chain planning: a case study of sales and operations planning. *Journal of Operations Management*, 29, 434-448.

Pauwels, K., Currim, I., Dekimpe, M.G., Ghysels, E., Hanssens, D.M., Mizik, N., & Naik, P. (2004). Modeling marketing dynamics by time series econometrics. *Marketing Letters*, 15, 167-183.

Ravichandran, T., & Liu, Y. (2011). Environmental factors, managerial processes, and information technology investment strategies. *Decisions Sciences,* 42, 537-574.

Rossana, R.J., & Seater, J.J. (1995). Temporal aggregation and economic time series. *Journal of Business & Economic Statistics*, 13, 441-451.

Williams, B.D., & Waller, M.A. (2010), Creating order forecasts: point-of-sale or order history? *Journal of Business Logistics*, 31, 231-251.

Yao, Y., & Dresner, M. (2008) "The inventory value of information sharing, continuous replenishment, and vendor-managed inventory." *Transportation Research Part E: Logistics and Transportation Review*, 44 (3), 361-378.

Yao, Y., Dong, Y., & Dresner, M. (2012), Supply chain learning and spillovers in vendor managed inventory, Decision Sciences, 43 (6), 979-992.

Zotteri, G., & Kalchschmidt, M. (2007). Forecasting practices: empirical evidence and a framework for research. *International Journal of Production Economics*, 108, 84-99.

Chapter 2 – LITERATURE REVIEW

II. LITERATURE REVIEW

A. THE BULLWHIP EFFECT

Causes of Bullwhip

The bullwhip effect is defined as “the amplification of demand variability from a downstream site to an upstream site” (Lee et al. 2004, p. 1887). It can result in substantial degrees of operations inefficiencies in supply chain management, thereby driving up costs to erode profits and service level (Lee et al., 1997). Due to increased variance, demand planners have to dedicate more production and transportation capacity, inventory, storage space, and other firm capital investments in anticipation of demand spikes, often paying premium prices (Haughton, 2009). When demand hits its deeper troughs, resources previously dedicated to fulfilling high demand simply sits idle. Moreover, rogue frequencies can induce pseudo-seasonality in demand patterns to mask the true market demand signal (Towill et al., 2007). Thus, bullwhip is a major contributor to inefficiencies, costs, and uncertainty within the supply chain that can result in lower levels of customer service (Lee et al., 1997). In order to mitigate the bullwhip effect and its associated costs, managers must accurately measure demand distortion in order to formulate appropriate responses.

The bullwhip effect gained substantial attention with Lee et al. (1997), which illustrated several causes of bullwhip based on managerial and behavioral idiosyncrasies at the retailer-supplier node of the supply chain. However, Forrester (1961) initially observed the existence of demand variance amplification in industrial dynamics. He identified that manufacturers in general seem to experience far greater demand fluctuations compared to lower echelons of the supply chain, particularly at the consumer level. Often, a 10% fluctuation in demand can be amplified to over 50% upstream along the supply chain. Based on this observation, he theorized that the observed amplification effect at every echelon of the supply chain is caused by the bounded rationality of managers. The principle reason is that the system complexity of multi-echelon demand exceeds the managerial intuition and capability required for its accurate assessment. As a result, managers responding to demand shocks cause amplified variance (Naish, 1994). Building on the “Forrester Effect”, a “beer distribution game” was developed (Sterman, 1989; Senge & Sterman, 1992; Paich & Sterman, 1993). In the game, participants emulate decision makers of a multi-echelon supply chain. The experiment demonstrated how small changes at the consumer level eventually developed into dramatic swings in demand at upper echelons of the supply chain. As managers attempt to avoid stock-outs, they tend to order more than is necessary (Kahn, 1987).

A separate line of theoretical cause of bullwhip is examined by Burbidge (1981). Burbidge elucidates that while demand can be amplified due to unintentional errors made by managers due to bounded rationality (i.e., Forrester Effect), rational managerial policies based on deliberate localized optimization can also amplify demand variability. Specifically, as managers attempt to reap scale economies associated with production and transportation system, they may intentionally produce in amounts greater than necessary to spread fixed costs across a higher number of units (Eichenbaum, 1989). Although the immediate financial impact is overall decreased average cost, increased quantities produced and transported result in longer time in between batches to heighten demand variability as well. Thus, increase in demand variability attributed to order-batching is known as “Burbidge Effect.”

Operationally, bullwhip can be amplified by shortened review intervals and increased minimum order size (Cachon, 1999), the adoption of continuous review policy (Chen and Samroengraja, 2004), and increased lead time variability (Chatfield et al., 2004) and the total cycle time of a supply chain (Gilbert, 2005). Behaviorally, Lee et al. (1997) identified several additional causes. First, demand signal is processed at each echelon of the supply chain. Due to both deliberate and unintentional actions, processed demand signals such as derived orders become distorted to amplify demand variance. Second, greater lead time induces uncertainty over time, so that demand forecast is less accurate. As a result, demand becomes distorted as buyers attempt to increase orders to quantities greater than true demand due to prior shortage. Alternatively, buyers may order at quantities less than true demand due to prior surplus. Third, during times of perceived shortage, buyers may deliberately order more under the belief that suppliers intend to only fulfill a fraction of demand, thereby corrupting true demand signal (Houlihan, 1987). And lastly, pricing variations tend to either increase or decrease short term demand, thereby further inducing demand volatility (Butman, 2003).

In addition to operational and behavioral causes, natural limitations of production technology may also result in the bullwhip effect. As materials and information flow through multiple echelons of the supply chain, delays between supply chain links contributes to increased variance due to less accurate forecasts (Blackburn, 1991). In general, communication and coordination are required to dampen the impact of bullwhip (Wu & Katok, 2006). To address bullwhip caused by information delays, systems such as electronic data interchange (EDI) (Disney & Towill, 2003; Machuca & Barajas, 2004) and vendor-managed inventory (VMI) (Waller et al., 1999) can simplify the decision hierarchy within the supply chain to improve information flow (Cantor & Katok, 2012). For physical distribution, substantial reduction to the bullwhip effect may be achieved through shortening lead times (Zhang, 2004) and controlled order-batching (Potter & Disney, 2006; Towill et al., 2007).

Empirical Evidence of Bullwhip

Since Forrester (1961) first noticed the tendency for demand variance to increase between industries, this phenomenon was observed in many industries at several levels (Table 1). At the industry level, Blanchard (1983) identified at the aggregate level that the American automobile industry exhibited strong evidence of bullwhip. Expanding upon this empirical observation, Blinder (1986), West (1986), and Krane and Braun (1991) all identified similar relationships between shipment and demand for various resource extraction, manufacturing, wholesaling, and manufacturing industries. More specifically, bullwhip can be readily observed in many industries, such as automotive (Blanchard, 1983; Lee et al., 1997; Cachon et al., 2007), computers and electronics (Blackburn 1991; Lee et al., 1997; Kaipia et al., 2006), both dry (Hammond, 1994; McKenney & Clark, 1995; Holstrom, 1997; Williams & Waller, 2010; 2011) and perishable groceries (Fransoo & Wouters, 2000; Lehtonen et al., 1999), personal care (Lee et al., 1997), and mechanical parts (McCullen & Towill, 2001).

Although overwhelming evidence supports the ubiquity of the bullwhip effect in supply chain, various studies indicated the opposite. Krane and Braun (1991) identified that approximately two-thirds of industries in their sample exhibited evidence of production smoothing. That is, the variance ratio of production and shipment to demand is less than one. More recently, empirical bullwhip literature also presented mixed findings on the presence and prevalence of bullwhip (Cachon et al., 2007; Bray & Mendelson, 2012; Chen & Lee, 2012). Cachon et al. (2007) concluded that the only retail sector to exhibit bullwhip is automobile. Dooley et al. (2009) attributes the lack of bullwhip in retail to the retailers' managerial use of inventory as a buffer to blunt the impact of demand uncertainty at the consumer level. However, when the level of analysis is taken from industry-monthly to firm-quarterly, Bray and Mendelson

(2012) concluded that over half of all firms in their sample showed bullwhip. Chen and Lee (2012) attribute this to the statistical masking effect of aggregation at both product-location and temporal dimensions.

B. DATA AGGREGATION

Forms and Rationale for Aggregation

Data aggregation can occur along two fronts. First, data can be aggregated across products and/or locations. For example, sales of single stock keeping unit (SKU) at multiple locations can be aggregated for predetermined purposes such as centralized distribution planning and forecasting. In inventory management, the reduction in system inventory due to managerial decisions to aggregate products based on location have been termed *portfolio effect* (Zinn et al., 1989). A principle statistical effect of product-location aggregation is risk-pooling (Gerchak & He, 2003). As long as demand histories for multiple products or multiple locations are not perfectly correlated, their variance of the aggregated series tends be less than the total variance of the two disaggregated series (Sucky, 2009).

Data can also be aggregated by time. Under temporal (time) aggregation, consecutive observations over time are aggregated into non-overlapping series at a greater time bucket. For example, most retailers make replenishment decisions on a weekly basis and therefore aggregate daily sales into weekly time buckets. Temporal aggregation’s statistical effect is similar to that of product-location aggregation. Stochastic variance over time can be reduced as the highs and lows offset each other (Hotta & Cardoso Neto, 1993). Similar to product-location aggregation, the degree of variance reduced in temporal aggregation is dependent on the nature and magnitude of a time series’ autocorrelation. As long as a time series is not perfectly positively autocorrelated, temporal aggregation will result in reduced variance.

Firms may choose to aggregate data for many reasons. First, different functions within each firm tend to measure performance at different levels and intervals. At the most basic level, a department within a retail establishment may keep track of its demand throughout the course of a day to replenish on-shelf inventory during less busy hours. However, the store as a whole may instead track demand in longer intervals to order from its warehouse or supplier. Alternatively, whereas a department manager may take keen interest in the highest performing individual SKUs, at the store level analysis tends to be performed based on category sales.

Second, whereas retail replenishment tends to operate on a weekly basis, marketing functions generally plan at the monthly level (Pauwels et al., 2004). Therefore, as cross-functional and inter-firm collaboration occurs, conflicts in levels of data aggregation tend to result. As industries move toward one-number forecasting (Finn, 2004), data aggregation becomes necessary to synchronize activities among functions and firms.

A third prominent reason for data aggregation to occur is to reduce data variance (Finn, 2004; Nikolopoulos et al., 2011). Due to statistical risk-pooling, data aggregation reduces the overall variance to assist firms with decisions such as inventory consolidation (Zinn et al., 1989; Ronen, 1990; Mahmoud, 1992; Evers, 1997; Ballou, 2005) and forecast efficiency (Hotta et al., 2005).

Negative Consequences of Aggregation

While variance reduction suggests reduction in volatility, it has substantial consequences in supply chain management. Variance in various logistics factors such as demand and lead time are principle determinants of inventory policies (e.g., safety stock). Both product-location and temporal aggregation may result in a masking effect on variance visibility (Chen & Lee, 2012). Further, aggregation can fundamentally transform a data series' statistical properties (Amemiya

& Wu, 1972; Rossana & Seater, 1995), thereby further complicating supply chain decisions that often rely on statistical results.

Data aggregation can have substantial effects the bullwhip effect. Mathematically, scholars caution that data aggregation can potentially result in masking the true degree of the bullwhip effect (Chen & Lee, 2012). Due to the statistical risk-pooling effect, in which highs and lows are offset as data become aggregated, the resultant variance ratio can present a distorted and overly-optimistic view of the magnitude of demand volatility to managers. In managing demand, efforts directed at controlling costs associated with demand volatilities should be based on demand patterns and peculiarities of each individual case (Disney et al., 2006). Therefore, planners that estimate aggregate bullwhip may see a biased level of variability.

Data aggregation can also affect statistical forecasting. For example, Williams and Waller (2011) identified that aggregation plays a definite role in the selection of demand signal input for forecasting. They find the risk-pooling property of product-location aggregation to be especially beneficial to suppliers in account-level demand planning. As the number of ship-to locations (e.g., DCs) increase, aggregation results in decreased stochastic variance to transform potentially volatile demand signals to become relatively stable. This benefit is expected to increase along with the number of locations being aggregated. While it is easy to assume that the risk-pooling property carries over to temporal aggregation to benefit suppliers as well, statistical properties of temporal aggregation presents a theoretical conflict. Namely, temporal aggregation can result in both information loss (e.g., Amemiya & Wu, 1972) and variance reduction (Hotta et al., 2005).

Three particularly transformative properties exist for temporal aggregation that can mislead demand planners. First, temporal aggregation can induce additional statistical properties such as moving-average residual structures [1](Amemiya & Wu, 1972; Brewer, 1973), two-way causality (Weiss, 1982), and previously unobserved error terms (Silvestrini & Veredas, 2008). Second, statistical properties at the disaggregate level may be lost or transformed, such as seasonality patterns (Wei, 1978; Drost & Nijman, 1993), autoregressive order [2](Stram & Wei, 1986), short-term cyclical variations (Rossana & Seater, 1995), and other general parameters (Silvestrini & Veredas, 2008). Both of these transformative effects can introduce substantial uncertainty to the forecast. However, a third statistical effect can potentially benefit forecast accuracy. Hotta et al. (2005) found that the variance reduction effect in data aggregation can result in the extreme highs and lows offsetting each other. As a result of reduced data variance, statistical forecast can be relatively more stable as well.

C. CUSTOMER DEMAND FORECASTING

Forecasting Practice

A prominent activity for demand planners is forecasting (Moon et al., 2003). Through the selective use of demand signals and forecasting model, a demand planner formulates a prediction for a future state of demand. With accurate predictions of future demand, planners may then allocate resources necessary to provide production, storage, transportation, and labor services necessary to fulfill anticipated orders. As a result, suppliers can determine the necessary

---

[1] Amemiya and Wu (1972) show that, for a variable that follows an AR model of order *p*, its aggregated variable also follows an AR model or order *p*, but with a MA residuals structure.

[2] Stram and Wei (1986) find that temporal aggregation can reduce the autoregressive order of ARIMA models.

inventory levels throughout its distribution network in anticipation of short term changes to demand (Mentzer & Cox, 1984).

Forecasting is considered an important firm function for various reasons. Accurate demand forecast allows firms to efficiently and effectively distribute resources to maximize service while minimizing costs. In particular, forecasting has substantial impact on the bullwhip effect. Naish (1994) argue that if accurate foreknowledge of demand changes are incorporated in to the planning process, demand volatilities can be successfully smoothed to determine production and capacity with greater ease. On the other hand, unreliable forecast contributes to downstream stockouts, which in turn may be met with over-ordering from buyers (Terwiesch et al., 2005). To increase forecast accuracy, scholars have developed myriad qualitative and quantitative forecasting methods designed to utilize demand signals amassed from many sources. Among most commonly utilized sources of forecast input are quantitative data such as historic transactions and macroeconomic variables.

Given today's volume of input data, demand forecasting is typically automated through a variety of software packages. Although the forecasting process had become substantially more user-friendly due to advances in software packages, the overall industry understanding of each method had not increased (McCarthy et al., 2004). While the vast majority of the most popular quantitative forecasting methods tend to be simplistic, users of forecasting packages continue to utilize a "black box" approach to forecasting (McCarthy et al., 2004, p. 322). That is, users do not always understand the various quantitative methods and simply assume that the software package always provides the optimal forecast.

Quantitative forecasting methods offered in today's software packages are highly diverse. Among all quantitative models, those based on exponential smoothing can provide forecast accuracy that rivals many other more complex models (Makridakis et al., 1982; Makridakis & Hibon, 2000). At the most basic level, users need to determine if a data series is classified as seasonal or non-seasonal (Chatfield & Yar, 1988). Subsequently, the appropriate forecasting model (e.g., seasonal or trend-adjusted) and its associated smoothing parameters must be specified a priori, such as trend, level, and seasonality. Although all smoothing parameters are continuous and may range between zero and one, a specific set of recommended ranges of value for each parameter is available (Silver et al., 1998). And finally, the appropriate demand signal will need to be selected for optimal demand forecast.

Forecasting in Retail

In the retail supply chain, large retailers generally depend on a network of regional distribution centers (DCs) to replenish stores. Retail stores both transmit aggregated point-of-sale (POS) and place orders to a single DC. In turn, the DC processes the store-level demand signal to place periodic orders to the supplier. To the supplier, orders from retail DCs signify their customer demand, which is different from demand signals at the market level (e.g., consumer demand). Therefore, suppliers most often rely on past DC orders to forecast future demand (Agarwal & Holt, 2005). However, retail orders tend to have amplified variance due to various causes (Lee et al., 1997). As a result, distorted demand can result in decreased forecast accuracy (Lee et al., 2000).

One way for demand planners to increase forecast accuracy is through incorporating shared information in to the forecasting process (Kiely, 1999; Romanow et al., 2004; Lapide, 1999; 2005). Recent advances to information technology had increasingly enabled real-time sharing of downstream demand signals. Because POS data more closely represents consumer demand, it is relatively free of the bullwhip effect. While companies such as Campbell Soup (Clark, 1994) and Barilla SpA (Hammond, 1994) reported substantial success through the use of shared consumer demand data, empirical evidence also indicates that the use of POS data frequently result in superior forecast accuracy for suppliers in the consumer-packaged goods (CPG) industry (Williams & Waller, 2010; 2011). Furthermore, this benefit increases along with logistics network complexity.

On the other hand, using POS data does not always lead to increased forecast accuracy. Since consumer demand signals are subject to retail managerial effects such as warehouse management systems (Autry et al., 2005), replenishment processes such as postponement (Zinn & Bowersox, 1988), and inventory management policies such as inventory consolidation (Evers & Beier, 1998), resulting patterns reflective of these factors are often embedded in retail orders. Therefore, if suppliers believe that future customer demand will likely follow past order patterns rather than POS demand, using DC orders as forecast input may yield superior forecast accuracy (Williams & Waller, 2010).

The importance of factoring retail ordering practice into forecasting is especially relevant to seasonal categories. For many industries, sales tend to occur in concentrated selling seasons (Fisher et al., 1994). As a result of anticipating demand spikes, retailers tend to build their inventories steadily during times of low demand, thereby avoiding potential capacity and inventory shortage (Cachon et al., 2007; Bray & Mendelson, 2012). This practice of seasonal smoothing often results in retail orders that tend to follow a “rhythmic” and predictable pattern that deviates from the consumer demand (Parkany, 1961). Thus, suppliers would more likely benefit from using past retail orders as indicators of future customer demand rather than incorporating demand signals from downstream along the supply chain.

Table 1 – Empirical literature on the presence of bullwhip among industries

Author	Industry	Product-Location	Temporal	Bullwhip Measure	BW/PS
Blanchard, 1983	Automotive	Brand / Division	Monthly	Sales variance ratio	BW
Blinder, 1986	Various	Industry	Monthly	Sales variance ratio	BW
West, 1986	Two-digit SIC	Sub-Industry	Monthly	Variance of change ratio	BW
Eichenbaum, 1989	Tobacco	Industry	Monthly	Unit variance ratio	BW
	Rubber	Industry	Monthly	Unit variance ratio	BW
	Food	Industry	Monthly	Unit variance ratio	BW
	Petroleum	Industry	Monthly	Unit variance ratio	BW
	Chemicals	Industry	Monthly	Unit variance ratio	BW
	Apparel	Industry	Monthly	Unit variance ratio	BW
Krane & Braun, 1991	Two-digit SIC	Sub-Industry	Monthly	Variance of change ratio	Both
Hammond, 1994	Dry Grocery	Firm, DC	Weekly	Unit variance ratio	BW
McKenney & Clark, 1995	Dry Grocery	Firm, Product	Not Specified	Unit variance ratio	BW
Leet et al. 1997a; 1997b	Automotive	Product	Not Specified	Unit variance ratio	BW
	Computers	Product	Not Specified	Unit variance ratio	BW
	CPG	Product	Not Specified	Unit variance ratio	BW

Lehtonen et al. 1999	Confectionery	Firm	Weekly	Coefficient of variation	BW
	Paper goods	Firm	Irregular	Coefficient of variation	BW
Fransoo & Wouters, 2000	Fresh Produce	Multiple	Multiple	Coefficient of variation	BW
	Refrigerated	Multiple	Multiple	Coefficient of variation	BW
McCullen & Towill, 2002	Mechanical parts	Firm	Non-specified	Coefficient of variation	BW
El-Beheiry et al. 2004	Toys	Firm	Weekly	Coefficient of variation	BW
Terwiesch et al. 2005	Semiconductor	Firm	Quarterly	Unit orders	BW
Kaipia et al., 2006	Electronics	Firm	Weekly	Unit comparison	BW
Cachon, Randall, & Schmidt, 2007	Retail	Industry	Monthly	Growth rate Variance	**PS**
	Wholesale	Industry	Monthly	Growth rate Variance	BW
	Manufacturing	Industry	Monthly	Growth rate Variance	Weak
Waller, Williams, & Eroglu, 2008	Retail	Product	½/4 Weeks	Unit variance	BW
Bray & Mendelson, 2012	Retail	Firm	Quarterly	Unit variance	BW
	Wholesale	Firm	Quarterly	Unit variance	BW
	Manufacturing	Firm	Quarterly	Unit variance	BW
	Resource Extraction	Firm	Quarterly	Unit variance	BW

D. REFERENCES

Agarwal, A., & Holt, G. (2005). Reducing inventory by simplifying forecasting and using point of sale data. *MLOG Thesis*, Massachusetts Institute of Technology.

Amemiya, T., & Wu., R.Y. (1975). The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association*, 67 (39), 628-632.

Allred, C.R., Fawcett, S.E., Wallin, C., & Magnan, G.M. (2012). A dynamic collaboration capability as a source of competitive advantage. *Decision Sciences*, 42, 129-161.

Ballou, R.H., (2005). Expressing inventory control policy in the turnover curve. *Journal of Business Logistics,* 26, 143-164.

Blackburn, J. D. (1991). The quick response movement in the apparel industry: A case study in time-compressing supply chains. In: Blackburn, J.D.(Ed.). *Time Based Competition: The Next Battleground in American Manufacturing*. Irwin, Homewood, IL (Chapter 11).

Blanchard, O. J. (1983). The Production and inventory behavior of the American automobile industry. *Journal of Political Economy*, 91 (3), 365-400.

Blinder, A.S. (1986). Can the production smoothing model of inventory behavior be saved? *The Quarterly Journal of Economics*, 101 (3), 431-454.

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

Brewer, K. R. W. (1973). Some consequences of temporal aggregation and systematic sampling for ARIMA and ARMAX models. *Journal of Econometrics*, 17 (1), 133-154.

Burbidge, J.L. (1981). The new approach to production. *Production Engineer*, 40 (12), 769-784.

Butman, J. (2003). A pain in the (supply) chain. *Harvard Business Review*, 80 (5), 31-44.

Cachon, G.P. (1999). Managing supply chain demand variability with scheduled ordering policies. *Management Science*, 45 (6), 843-856.

Cachon, G. P., Randall, T., & Schmidt. G.M. (2007). In search of the bullwhip effect. *Manufacturing and Service Operations Management*, 9 (4), 457-479.

Cantor, D.E., & Katok, E. (2012). Production smoothing in a serial supply chain: A laboratory investigation. *Transportation Research Part E*, 48, 781-794.

Chatfield, C. & Yar, M. (1988). Holt-Winters forecasting: some practical issues. *Journal of the Royal Statistical Society*, 37 (2). 129-140.

Chatfield, D.C., Kim, J.G., Harrison, T.P., & Hayya, J.C. (2004). The bullwhip effect—Impact of stochastic lead time, information quality, and information sharing: a simulation study. *Production and Operations Management*, 13 (4), 340-353.

Chen, F., & Samroengraja, R. (2004). Order volatility and supply chain costs. *Operations Research*, 52 (5), 707-722.

Chen, L., & Lee, H.L. (2012). Bullwhip effect measurement and its implications. *Operations Research*, 60, 771-784.

Clark, T. (1994). Campbell soup: A leader in continuous replenishment innovations. *Harvard Business School Case*, Boston, Massachusetts.

Disney, S.M., & Towill, D.R. (2003). The effect of VMI dynamics on the bullwhip effect in supply chains. *International Journal of Production Economics*, 85, 199-215.

Disney, S.M., Farasyn, I., Lambrecht, M., Towill, D.R., & Van de Velde, W. (2006). Taming the bullwhip effect whilst watching customer service in a single supply chain echelon. *European Journal of Operational Research*, 173, 151-172.

Drost, F.C., & Nijman, T.E. (1993). Temporal aggregation of GARCH processes. *Econometrica*, 61 (4), 909-927.

Eichenbaum, M.S. (1989). Some empirical evidence of the production level and production cost smoothing models of inventory investment. *American Economics Review*, 79 (4), 853-864.

El-Beheiry, M., Wong, C.Y., & El-Kharbotly, A. (2004). Empirical quantification of bullwhip effect (with application to a toy supply chain). *Proceedings of the 13th International Working Seminar on Production Economics*, Igls, Austria, 3 (February), 83-95.

Evers, P.T. (1997). Hidden benefits of emergency transshipments. *Journal of Business Logistics*, 18, 55-76.

Evers, P.T., & Beier, F.J. (1998). Operational aspects of inventory consolidation decision making. *Journal of Business Logistics*, 19, 173-189.

Finn, R. (2004). Weekly vs monthly forecasting in the supply chain. *Logistics and Transport Focus*, 6, 22-28.

Fisher, M., Hammond, J.G., Obermeyer, W.R., & Raman, A. (1994). Making supply meet demand in an uncertain world. *Harvard Business Review*, 72 (3), 83-93.

Forrester, J.W. (1961). *Industrial Dynamics*. Productivity Press, Portland OR.

Fransoo, J.C., & Wouters, M.J.F. (2000). Measuring the bullwhip effect in the supply chain. *Supply Chain Management: An International Journal*, 5(2), 78-89.

Gerchak, Y., & He, Q.M. (2003). On the relation between the benefits of risk pooling and the variability of demand. *IIE Transactions*, 35 (11), 1027-1031.

Hammond, J. (1994). Barilla SpA (a). HBS Case 9-694-046, Harvard Business School, Boston, MA.

Haughton, M.A. (2009). Distortional bullwhip effects on carriers. *Transportation Research Part E*, 45, 172-185.

Homlstrom, J. (1997). Product range management: A case study of supply chain operations in the European grocery industry. *International Journal of Supply Chain Management*. 2 (8), 107-115.

Hotta, L.K., & Cardosc Neto, J. (1993). The effect of aggregation on prediction in autoregressive integrated moving-average models. *Journal of Time Series Analysis,* 14 (3), 261-269.

Hotta, L.K., Valls Pereira, P.L., & Otta, R. (2005). Effect of outliers on forecasting temporally aggregated flow variables. *Sociedad de Estadistica e Investigacion Operativa Test*, 13, 371-402.

Houlihan, J.B. (1987). International supply chain management. *International Journal of Physical Distribution and Materials Management*, 17 (2), 51-66.

Kahn, J.A. (1987). Inventories and the volatility of production. *The American Economic Review,* 77 (4), 667-679.

Kaipia., R., Korhonen, H., & Hartiala, H. (2006). Planning nervousness in a demand supply network: An empirical study. *International Journal of Logistics Management*, 17 (1), 95-113.

Kiely, D. (1999). Synchronizing supply chain operations with consumer demand using customer data. *Journal of Business Forecasting*, 17 (4), 3-9.

Krane, S.D., & Braun, S.N. (1991). Production smoothing evidence from physical-product data. *Journal of Political Economy*, 99 (3), 558-581.

Lapide, L. (2005). Account level forecasting needs downstream data. *Journal of Business Forecasting*, 24 (2), 17-19.

Lapide, L. (1998). New developments in business forecasting. *Journal of Business Forecasting Methods & Systems,* 17 (4). 29-30.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43 (4), 546-559.

Lee, H.L., So, K.C., & Tang, C.S. (2000). The value of information sharing in a two-level supply chain. *Management Science,* 46 (5). 626-643.

Lee, H.L., Padmanabhan, V., & Whang, S. (2004). Comments on 'information distortion in a supply chain: The bullwhip effect'—The bullwhip effect: Reflections. *Management Science*, 50 (12), 1887-1893.

Lehtonen, J., Holmstrom, J., & Slotte, J. (1999). Constraints to quick response systems in the implosive industries. *Supply Chain Management: An International Journal*, 4 (1), 51-57.

Machuca, J.A., & Barajas, R.P. (2004). The impact of electronic data interchange on reducing bullwhip effect and supply chain inventory costs. *Transportation Research Part E*, 40, 209-228.

Mahmoud, M.M. (1992). Optimal inventory consolidation schemes: a portfolio effect analysis. *Journal of Business Logistics*, 13, 193-214.

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: results of a forecasting competition. *Journal of Forecasting,* 1, 111-153.

Makridakis, S., & Hibon, M. (2000). The M3-competition: results, conclusions and implications. International *Journal of Forecasting,* 16, 451-476.

McCarthy, T.M., Davis, D.F., Golicic, S.L., & Mentzer, J.T. (2006). The evolution of sales forecasting management: a 20-year longitudinal study of forecasting practices. *Journal of Forecasting*, 25 (5). 303-324.

McCullen, P., Towill, D.R. (2002). Diagnosis and reduction of bullwhip in supply chains. *Supply Chain Management: An International Journal*, 7 (3), 164-179.

McKenney, J. L., & Clark, T.H. (1995). Campbell Soup Co.: A leader in continuous replenishment innovations, HBS Case 9-195-124, Harvard Business School, Boston, MA.

Mentzer, J.T., & Cox, J.E. Jr. (1984). Familiarity, application, and performance of sales forecasting techniques. *Journal of Forecasting*, 14, 465-476.

Moon, M.A., Mentzer, J.T., & Smith, C.D. (2003). Conducting a sales forecasting audit. *International Journal of Forecasting*, 19, 5-25.

Naish, H.F. (1994). Production smoothing in the linear quadratic inventory model. *Quarterly Journal of Economics*, 104, 864-875.

Nikolopoulos, K., Syntetos, A.A., Boylan, J.E., Petropoulous, F., & Assimakopoulos, V. (2011). An aggregate-disaggregate intermittent demand approach (ADIDA) to forecasting: an empirical proposition and analysis. *Journal of the Operational Research Society*, 62, 544-554.

Paich, M., & Sterman, J.D. (1993). Boom, bust, and failures to learn in experimental markets. *Management Science*, 39 (12), 1439-1458.

Parkany, J. (1961), A new approach to sales forecasting and production scheduling. *Journal of Marketing,* 25 (3), 14-21.

Pauwels, K., Currim, I., Dekimpe, M.G., Ghysels, E., Hanssens, D.M., Mizik, N., & Naik, P. (2004). Modeling marketing dynamics by time series econometrics. *Marketing Letters*, 15, 167-183.

Potter, A., & Disney, S.M. (2006). Bullwhip and batching: an exploration. *International Journal of Production Economics*, 104 (2), 408-418.

Romanow, K., Suleski, J., Hofman, D., Kirby, P., & Preslan, L. (2004). POS data: The beginning of DDSN for consumer products manufacturers. Retrieved from www.amrresearch.com

Ronen, D. (1990). Inventory centralization: the square root law revisited again. *Journal of Business Logistics*, 11, 129-138.

Rossana, R.J., & Seater, J.J. (1995). Temporal aggregation and economic time series. *Journal of Business & Economic Statistics*, 13, 441-451.

Senge, P.M., & Sterman, J.D. (1992). Systems thinking and organizational learning: Acting locally and think globally in the organization of the future. *European Journal of Operational Research*, 59 (1), 137-150.

Silver, E.A., Pyke, D.F., & Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling*, 3rd ed., New York: John Wiley & Sons.

Silvestrini, A., & Veredas, D. (2008). Temporal aggregation of univariate and multivariate time series models: a survey. *Journal of Economic Surveys*, 22, 458-497.

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

Stram, D.O., & Wei, W.W.S. (1986), A methodological note on the disaggregation of time series totals. *Journal of Time Series Analysis*, 7 (4), 293-302.

Sucky, E. (2009). The bullwhip effect in supply chains—an overestimated problem? *International Journal of Production Economics*, 118, 311-322.

Terwiesch, C., Ren, T., Ho, H., & Cohen, M. (2005). Forecast sharing in the semiconductor equipment supply chain. *Management Science*, 51, 208-220.

Towill, D.R., Zhou, L., & Disney, S.M. (2007). Reducing the bullwhip effect: looking through the appropriate lens. *International Journal of Production Economics*, 108, 444-453.

Waller, M., Johnson, M.E., & Davis, T. (1999). Vendor-managed inventory in the retail supply chain. *Journal of Business Logistics*, 20 (1), 183-203.

Wei, W.W.S. (1978). Some consequencoes of temporal aggregation in seasonal time series models. *Seasonal Analysis of Economic Time Series*, Washington D.C., U.S. Department of Commerce, Bureau of the Census.

Weiss, A.A. (1984). Systematic sampling and temporal aggregation in time series models. *Journal of Econometrics,* 26 (3), 271-281.

West, K.D. (1986). A variance bounds test of the linear quadratic inventory model. *Journal of Political Economy*, 94 (4), 374-401.

Williams, B.D., & Waller, M.A. (2010). Creating order forecasts: point-of-sale or order history? *Journal of Business Logistics*, 31, 231-251.

Williams, B.D., & Waller, M.A. (2011). Top-down vs. bottom-up demand forecasts: the value of shared point-of-sale data in the retail supply chain. *Journal of Business Logistics,* 32, 17-26.

Wu, Y., & Katok, E. (2006). Learning, communication, and the bullwhip effect. *Journal of Operations Management*, 24 (6), 839-850.

Zhang, X. (2004). The impact of forecasting methods on the bullwhip effect. *International Journal of Production Economics,* 88, 15-27.

Zinn, W., & Bowersox, D.J. (1988). Planning physical distribution with the principle of postponement. *Journal of Business Logistics*, 9, 117-136.

Zinn, W., Levy, M., & Bowersox, D.J. (1989). Measuring the effect of inventory centralization/decentralization on aggregate safety stock: the ‘square root law’ revisited. *Journal of Business Logistics,* 10, 1-14.

Chapter 3 – THE EFFECT OF DATA AGGREGATION ON BULLWHIP

III. THE EFFECT OF DATA AGGREGATION ON BULLWHIP

A. INTRODUCTION

The bullwhip effect is defined by Lee et al. (1997) 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)” (p.546). Empirical research has found evidence of the bullwhip effect in many industries, including automotive (Blanchard, 1983; Blackburn, 1991; Lee et al. 1997; Cachon et al. 2007), apparel (Blackburn 1991), computers and electronics (Blackburn 1991; Lee et al. 1997; Kaipia, Korhonen, & Hartiala, 2006), both dry (Hammond, 1994; McKenney and Clark, 1995; Holstrom, 1997) and perishable groceries (Fransoo & Wouters, 2000), personal care (Lee et al. 1997), and mechanical parts (McCullen & Towill, 2002). Recent empirical studies have called these findings into question. In particular, Cachon et al. (2007) found that bullwhip is largely absent in many industries, most notably the retail industry. In fact, five of the six retail industries studied by these authors did not exhibit bullwhip, and the authors conclude that, in general, retailers are production smoothers. Production smoothing, developed in the economics literature, explains that a firm can smooth its production relative to sales by buffering with inventory. To further the empirical investigation for the bullwhip effect, Bray and Mendelson (2012) studied bullwhip at the firm level and found mixed results within retail segments. That is, some segments exhibited bullwhip behavior while others indicated a proclivity to smooth.

While literature yields inconsistent conclusions regarding the presence and prevalence of bullwhip in the retail industry, both Cachon et al (2007) and Bray and Mendelson (2012) cite data aggregation as a potential confounding factor to their findings. Cachon et al. (2007) utilized quarterly, industry-level data and indicated that bullwhip may be more prevalent if measured at lower levels of aggregation. Bray and Mendelson (2012) elucidated a need to use high frequency product-level order and demand data to gain further understanding of bullwhip's existence, rather than firm- and industry-level data.

In this research, we claim that whether retailers bullwhip or production smooth may have significant planning implications. Thus, our primary objective is to empirically test the effect of data aggregation on the observation of a retailer's bullwhip measurement. We operationalize this research by using retailer sales and order data taken from three consumer product categories, obtained at the product-weekly level. We follow Bray and Mendelson (2012) and Cachon et al. (2007) and classify bullwhip as when the demand variability at an upstream echelon (e.g., order) is greater than at a downstream echelon (e.g., sales). In addition, we not only calculate bullwhip by log differencing (Cachon et al. 2007) and with coefficient of variation (Fransoo & Wouters 2000), we also calculate it as the unit variance ratio as originally conceptualized by Lee at al. (1997) and utilized by Bray and Mendelson (2012). By examining data aggregation's effect on the retailer's bullwhip measurement, we are also able to discuss the effect that the different measurement methods may have on findings throughout the extant literature .

Chen and Lee (2012) recently analyzed data aggregation's influence on bullwhip under assumptions of specific demand processes, seasonality, and spatial independence. Using actual sales data and an empirical model that accounts for the various assumptions above, we find that both product-location (e.g., product to category) and temporal (e.g., week-to-month) aggregation have a significant masking effect on bullwhip measurement.

In the following section, we briefly discuss how bullwhip is measured. Subsequently in §3, we briefly discuss the literature and conceptual basis of our study. We then present in §4 details of our data. In §5, we take a cursory look at the nature of the amplification ratio as it is compared across the three methods of calculation, and proceed to discuss our method of analysis to examine the effect of empirical data aggregation on the amplification ratio. Finally, we present our conclusions.

B. BULLWHIP MEASUREMENT

We begin by considering a batch-order inventory system composed of a set of retail stores which are fulfilled by a single retail distribution centers (DC). Store-level sales information for product I is transmitted to DC k, whose orders are fulfilled by the supplier. Let O_{ikt} be the quantity ordered of product I by retailer DC k in period t and D_{ikt} be the sum of point-of-sale (POS) data at time t of product I for all stores served by DC k. Thus order receipts and POS demand variance may be matched by product I, DC k, over a period of time t to create measures of bullwhip. The current empirical bullwhip literature primarily measures the variance ratio in three ways. Cachon et al. (2007) estimates bullwhip with the fractional growth rate method. Bray and Mendelson (2012) utilizes the coefficient of variation method (c.f., Fransoo & Wouters 2000). Lastly, Chen and Lee (2012) uses the unit variance ratio method as originally conceptualized in Lee et al. (1997).

Using the fractional growth rate measure denoted as $Growth$, (Cachon et al. 2007), sales and order variance are estimated by first differencing the natural log of both series. This approach detrends the data in order to account for the time-dependent portion of total variance (c.f., Cachon et al. 2007, p. 463, footnotes 6, 7, & 8).

$$Growth = \frac{V[Ln(O_{ikt}) - Ln(O_{ikt-1})]}{V[Ln(D_{ikt}) - Ln(D_{ikt-1})]}$$

Next, we review the coefficient of variation measure ($\boldsymbol{CV}$). This measure (e.g., Fransoo & Wouters 2000) and other variations of it (e.g., Dooley et al. 2009) are commonly utilized when data does not contain explicit buyer-seller relationships (c.f., Bray & Mendelson 2012, p. 2) and is also among the most popular measures used in the empirical bullwhip literature (e.g., Fransoo and Wouters 2000; McCullen and Towill 2002; El-Beheiry et al. 2004).

$$CV = \frac{S(O_{ikt}) / \sum_{t=1}^{T}(O_{ikt}) * \frac{1}{T}}{S(D_{ikt}) / \sum_{t=1}^{T}(D_{ikt}) * \frac{1}{T}}$$

Finally, we review the unit variance (Var) approach to bullwhip measurement (e.g., Chen & Lee 2012; Torres and Maltz 2008; Dejonckheere et al. 2003; Lee et al. 1997; Sterman 1989). Although the unit variance measure is similar to the coefficient of variation measure, the following distinctions should be noted: (1) variance and standard deviation are not collinear; and (2) the coefficient of variation ratio is a normalized measure, which may further reduce the degree of correlation between these two measures. Further, the supply chain bullwhip effect is described consistently as an increase in total variability from sales to the buyer to orders to the supplier (e.g., Lee et al. 1997; Cachon et al. 2007; Bray & Mendelson 2012). Therefore our third amplification ratio is measured as:

$$Var = \frac{V(O_{ikt})}{V(D_{ikt})}$$

C. AGGREGATION AND SEASONALITY

Aggregation Effect

Data aggregation results in lower aggregated variance as extreme highs and lows are offset (Amemiya & Wu 1975; Rossana & Seater 1995). Utilizing the unit variance measure, Chen and Lee (2012) analytically show that product and location aggregation results in a smoothing effect to mask the bullwhip effect due to batch-ordering. That is, data aggregation leads to variance reduction which may disguise the magnitude of the bullwhip effect. In fact, under assumptions of no capacity limit, no batch-ordering, and spatial independence, aggregation of products with common seasonality or an AR(1) demand process results in a monotonically decreasing unit variance ratio (Chen & Lee 2012). For a complete derivation, we refer readers to propositions 8 & 9 of Chen & Lee (2012). That is, product-location aggregation decreases observed bullwhip for products with common seasonality and AR(1) demand process.

In addition to product-location aggregation, temporal aggregation may exhibit a similar masking effect (Chen & Lee 2012). In the econometrics literature, studies have shown that temporal aggregation results in loss of variance when data is the result of an autoregressive process (Amemiya & Wu 1972), autoregressive moving average models with exogenous variables (Brewer 1973), seasonal structures (Wei 1978) and nonstationary models (Tiao 1972). Under the assumption of no capacity limit and no batch-ordering, with a constant lead time and an ARMA(1,1) demand process, Chen and Lee (2012) show that the unit variance ratio also monotonically decreases toward unity as temporal aggregation increases. We refer readers to proposition 7 of Chen and Lee (2012) for a more detailed explanation of this effect. Therefore like product-location aggregation, temporal aggregation also results in a "masking" effect of observed bullwhip.

Seasonal Effect

The literature recognizes that demand variance is generally composed of seasonal and stochastic components (e.g., Sobel 1969). While seasonal variance is due to recognizable and predictable demand patterns, stochastic variance is due to randomness. As Bray and Mendelson (2012) argue, firms competing in the same industry are likely to be affected by similar seasonal signals. Therefore as product-location aggregation occurs, seasonal variance is preserved due to highly correlated seasonal effects while firm-specific (in our case, product-specific) shocks are attenuated.

In propositions 2 and 3, Chen and Lee (2012) demonstrate the inclusion of seasonality in bullwhip measurement will have a stabilizing effect on observed bullwhip. Indeed, the extant empirical literature seems to confirm this notion. Ghali (1987) observes lowered upstream demand variance when using seasonally unadjusted data from the cement industry. In a more comprehensive examination at the industry level, Cachon et al. (2007) find similar results also using seasonally unadjusted data. Chen and Lee (2012) further demonstrate that when there is no capacity limit or batch-ordering, the bullwhip ratio tends to decrease as the variance component of seasonality dominates the stochastic component. In decomposing seasonality and stochastic variance components, Bray and Mendelson (2012) conclude that whereas firms' ordering behavior can smooth the predictable seasonal variations, they instead amplify stochastic shocks. Further, the reduction in variance amplification due to seasonality is different from reduction due to aggregation. Whereas aggregation reduces the bullwhip ratio through a "masking" effect, seasonality may induce smoothing of orders, leading to dampended bullwhip measures. Therefore we expect to observe that those products with higher seasonal variance will likely to exhibit lower observed bullwhip.

Controls

In addition to aggregation and seasonal effects described above, autocorrelated demand can potentially influence bullwhip measurement (Cachon et al. 2007). Kahn (1987) 42ehavior42s that a positive association exists between the autocorrelation in demand and observed bullwhip. Therefore, like Cachon et al. (2007), we account for potential effects of autocorrelation to more accurately assess the effects of aggregation and seasonality on bullwhip measurement.

D. ESTIMATION PROCEDURE

Data Collection

To conduct our analyses, we collected data from a consumer packaged goods (CPG) manufacturer which produces and markets a wide variety of product lines. The sample contains both sales and order data. The data was collected on a weekly basis across 10 customer distribution centers (DC) which are located throughout the United States for 9 non-seasonal, dry grocery products, 7 non-seasonal, perishable products and 7 seasonal, dry grocery products over a two-year period. In Table 1, we compare and contrast these product categories across several dimensions.

Variable Specification

We begin by pairing sales and order data at the product-weekly level. We then aggregate products based on their respective categories to create category-weekly series. A binary variable $PAgg$ is used to denote product-location aggregation where the product-level and category-level observations are coded as $PAgg \in [0,1]$, respectively. Further, we temporally aggregate the product and category series from the weekly to the monthly level. Thus, a second binary variable $TAgg$ is created to denote temporal aggregation where the weekly-level and monthly-level observations are coded as $TAgg \in [0,1]$, respectively.

To account for seasonality we use sales data to derive the variance of the underlying seasonal index (*SEAS)* for weekly unit sales over the two-year period at each DC. To calculate the seasonal index (Gaynor & Kirkpatrick 1993), let D_{ikt} denote sales for product *I* at DC *k* at time *t*, $SEAS_{ikt}$ is thus calculated as the variance of the ratio of each observation to the average over $T = 104$ periods:

$$SEAS_{ikt} = V\left(\frac{D_{ikt}}{\sum_{t=1}^{T}(D_{ikt}) * \frac{1}{T}}\right)$$

In addition to seasonality, we also obtain the first order autocorrelation factor (*AF)* for each product for their weekly unit sales at each DC through a simple AR(1) regression for each product *I* and DC *k* over the two-year period (*t* = 104):

$$D_{ikt} = \varphi D_{ikt-1} + \varepsilon_t$$

The estimated coefficient φ is recorded as a continuous variable.

E. ANALYSIS

Observed Bullwhip

All three bullwhip measures calculated at the weekly and monthly levels exhibit bullwhip behavior for both the product and category levels. We report these means in Table 2. We first compare the average bullwhip for the three measures at the product level. For $Growth$, the average bullwhip ratio decreased from 331.963 at the weekly level to 17.447 at the monthly level. For CV and Var, the observed decreases are from 3.117 to 1.713 and 12.913 to 3.780, respectively. At the category level, similar decreases are observed for the three measures. They are 64.820 to 6.357, 2.397 to 1.540, and 7.250 to 3.130, for $Growth$, CV, and Var, respectively. Further, we observe decreased bullwhip ratios for all three measures from the product level to the category level at both weekly and monthly levels of temporal aggregation as well. All differences are statistically significant at the 0.01 level.

While examining the bullwhip means offers potential support for both aggregation effects, seasonality's effect on bullwhip warrants a more detailed examination. Table 3 reports the estimated bullwhip means at both product and category levels and at both the weekly and monthly levels, for all three measurement methods. Recall that we have three distinct categories, each with its own set of characteristics (seasonality, product life cycle, and shelf-life). Similar to Table 2, we observe a similar pattern in measured bullwhip, where both temporal and product-location aggregation decrease bullwhip.

Comparing the bullwhip measures reveal several noticeable differences. First, it appears that the fractional growth rate measure results in higher observed bullwhip at the lowest levels of aggregation. Second, though the seasonal product category tends to yield the lowest bullwhip for both the coefficient of variation and the unit variance measures, it actually yields the highest when estimated with fractional growth rate. Third, evaluating all minimums, means and maximums indicate that product-location aggregation has a "compression" effect which raises the minimum bullwhip and substantially lowers the maximum bullwhip, while simultaneously reducing the mean bullwhip. Further, the bullwhip means for all three measures and all product-location and temporal aggregation combinations suggest bullwhip is present.

In terms of the prevalence of bullwhip in the retail industry, Cachon et al. (2007) found bullwhip only in the automotive retail segment at the industry level of analysis. At the firm level, Bray and Mendelson (2012) found bullwhip in general merchandise, furniture, and other non-categorized retailers. In our sample, we find bullwhip for all product-location and temporal aggregation levels for the fractional growth rate measure. For the coefficient of variation bullwhip measure, we find negative bullwhips for 6.5% and 3.3% of the product-monthly and category-monthly observations, respectively. For the unit variance measure we find negative bullwhips for 20.28% and 6.67% of the product-monthly and category-monthly observations, respectively.

Estimating the Masking and Dampening Effects on Bullwhip Measurement

In our empirical investigation of the effects of data aggregation on bullwhip measurement, we have three dependent variables of interest, one for each method of measurement ($Growth_{i,k,t}, CV_{i,k,t}, Var_{i,k,t}$). As Bray and Mendelson (2012, p.15) indicate that "the effect [is] idiosyncratic, as the bullwhip varies greatly across firms." Instead of the firm level, our product-DC observations are nested in category and DC levels. We expect that there are idiosyncratic effects across both categories and DCs, and thus, to account for these unobservable effects, we use hierarchical linear modeling (HLM). HLM allows observations to be nested within higher-level categories and accounts for the lack of independence among observations due to the multi-level structure of the data (Raudenbush & Bryk 2002). For example, observations of products taken from the same category or a particular DC cannot be assumed to be independent, thus HLM allows variance to be parceled out at these higher-level structures. We estimate the model using full maximum likelihood similar to Ang et al. (2002), DeHoratius and Raman (2008), and Liao and Chuang (2004).

Null Model

First, we partition the dependent variables into the variance across products *I* (*I* = 1,…,26),categories *j* (*j*=1,2,3), DCs *k* (*k*=1,…,10), and temporal aggregation levels *m* (m = 0,1), where 1 indicates temporal aggregation to the monthly level. We specify the null model as,

$$Growth_{ijkm} = \theta_0 + CAT_{00jm} + DC_{00km} + e_{ijkm},$$

$$CV_{ijkm} = \theta_0 + CAT_{00jm} + DC_{00km} + e_{ijkm},$$

$$Var_{ijkm} = \theta_0 + CAT_{00jm} + DC_{00km} + e_{ijkm},$$

where Θ_0 is the fixed intercept parameter, while the random effect parameter of category j is CAT_{00jm} and the random effect parameter of DC k is DC_{00km}. Finally, the random effect parameter of product I is e_{ijkm}. Note that T_{00jm} , DC_{00km} and e_{ijkm} are each normally distributed with a zero mean and variances of τ_{CAT00}, τ_{DC00}, and σ^2, respectively.

<u>Conditional Model</u>

We then include our independent variables of interest in the model. $TAgg_{ijkm}$, $PAgg_{ijkm}$, $SEAS_{ijkm}$, and AF_{ijkm} are the predictor variables. The first two variables test the effect of temporal and product-location aggregation on the bullwhip measures as proposed by Chen and Lee (2012). The latter two variables are included to examine non-aggregation influences on the bullwhip measurements. In this model design, we assume that aggregation's effect is fixed across categories and DCs, rather than randomly varying (Raudenbush & Bryk 2002). In addition, we assess model significance by examining the difference in negative log-likelihood between models and report the χ^2value of model change, as well as the associated statistical significance. Our full model is specified as,

$$Growth_{ijkm} = \theta_0 + CAT_{00jm} + DC_{00km} + e_{ijkm}$$
$$+\beta_{1,Growth} * TAgg_{ijkm} + \beta_{2,Growth} * PAgg_{0jkm} + \beta_{3,Growth} * SEAS_{ijkm} + \beta_{4,Growth} * AF_{ijkm},$$

$$CV_{ijkm} = \theta_0 + CAT_{00jm} + DC_{00km} + e_{ijkm}$$
$$+\beta_{1,CV} * TAgg_{ijkm} + \beta_{2,CV} * PAgg_{0jkm} + \beta_{3,CV} * SEAS_{ijkm} + \beta_{4,CV} * AF_{ijkm} +,$$

$$Var_{ijkm} = \theta_0 + CAT_{00jm} + DC_{00km} + e_{ijkm}$$

$$+\beta_{1,Var} * TAgg_{ijkm} + \beta_{2,Var} * PAgg_{0jkm} + \beta_{3,Var} * SEAS_{ijkm} +$$

$$\beta_{4,Var} * AF_{ijkm} +,$$

where β_1 and β_2 are the fixed effects for temporal and product-location aggregation levels, $TAgg_{ijkm}$ and $PAgg_{0jkm}$, respectively. B_3 is the estimated fixed effect for each product's seasonality factor $SEAS_{ijkm}$ at the weekly-DC-product level. And finally, β_4 is the estimated fixed effect for the first order autocorrelation coefficient for each product-DC series over the 104 week period, AF_{ijkm}.

Prior to model estimation, we transformed all dependent variables and $Seas_{ijkm}$into their natural log form. This transformation provides two benefits. Analysis of the raw data suggests some nonlinearity in the relationships between the outcome variable and the predictors. The natural log transformation process thus induces linearity in the regression model (Kleinbaum et al. 1998; DeHoratius & Raman, 2008). In addition, the transformation also addresses the potential skewness in our dependent variables.

F. HLM RESULTS

Overall, the conditional models for all three dependent variables demonstrate superior fit to their respective null models, as indicated by the significant χ^2 values for all three measures. In addition, the fixed effects in our conditional models explain 52.09%, 56.86%, and 60.03% of between-product variances for $Growth_{ijkm}$, CV_{ijkm} and Var_{ijkm}, respectively. From the estimation results (Null Models, Table 4), we find that 91.97% of the variance for $Growth_{ijkm}$ exists across products and 8.03% across product categories, with no significant variance across DCs. In addition, we find that 50.25%, 40.89%, and 8.87% of the variance for CV_{ijkm} exists across products, product categories, and DCs, respectively. The variance components for

Var_{ijkm} are 56.31%, 36.98%, and 6.71% across products, product categories, and DCs. Accounting for significant between-category and between-DC effects allows us to test whether the residual variance for our three dependent measures is associated with product-level measurements and aggregation effects. In the following discussion, aggregation's effect is fixed for all products at the DC level for CV_{ijkm} and Var_{ijkm}, and category level for $Growth_{ijkm}$.

Chen and Lee (2012) propose that product-location aggregation and temporal aggregation both result in dampened observation of bullwhip measurement, given that a particular set of assumptions are true. Based on Chen and Lee's propositions (2012), we systematically aggregate product-weekly observations to category-monthly and test aggregation's effect on the bullwhip ratio based on three common industry measures.

Across all three measures, we find that temporal ($TAgg_{ijkm}$) and product-location ($PAgg_{ijkm}$) aggregation mask bullwhip observation. Temporally, aggregating weekly demand information to the monthly level resulted in negative effects on the bullwhip measurements ($\beta_{1,Growth} = -2.107, p < 0.01$; $\beta_{1,CV} = -0.410, p < 0.01$; $\beta_{1,Var} = -1.087, p < 0.01$). For product-location aggregation, product-level observations and category-level observations are compared through the parameter estimate for *Pagg*. We find the parameter estimate to be negative and significant ($\beta_{2,Growth} = -1.339, p < 0.01$; $\beta_{2,CV} = -0.260, p < 0.05$; $\beta_{2,Var} = -0.439, p < 0.05$).

Chen and Lee (2012) also proposed that the inclusion of seasonality results in dampened bullwhip measurement. We find that $Seas_{ijkm}$ also significantly lowers bullwhip measurement for all three measures ($\beta_{3,Growth} = -0.258, p < 0.01$; $\beta_{3,CV} = -0.239, p < 0.01$; $\beta_{3,Var} = -0.529, p < 0.01$). As Bray and Mendelson (2012) explain, a greater seasonal component of total variance leads to a dampening effect. Chen and Lee (2012) also state that including seasonality in bullwhip measurement tends to result in a stabilizing effect. Further, these results are also consistent with Cachon et al. (2007)'s observation that when the amplification ratio is greater than one, the ratio measurement including seasonality is less than that excluding seasonality in most cases.

Unlike Cachon et al. (2007), we find significant positive association between the autoregressive coefficient and the fractional growth rate measure ($\beta_{4,Growth} = 2.086, p < 0.01$) but not the coefficient of variation ($\beta_{4,CV} = 0.031, p > 0.10$) and unit variance ($\beta_{4,Var} = -0.223, p > 0.10$) measures.

G. CONCLUSIONS

Cachon et al. (2007) conclude that retailers generally do not bullwhip but instead smooth demand. This finding was later disputed by Bray and Mendelson (2012) through firm-level, rather than industry-level analyses. Chen and Lee (2012) contend that the observed discrepancy is likely caused by the effects of data aggregation and the dominance of seasonal variance. The current research contributes to the literature by utilizing product-weekly level data to empirically examine the effect of product-location and temporal aggregation and seasonality on bullwhip measurement in a retail context. Our results suggest that bullwhip measurement may in fact be masked by data aggregation effects and dampened by seasonality, yielding potentially confounding results when measuring bullwhip at the industry or firm level.

These empirical findings are particularly salient to retail suppliers, such as consumer packaged goods (CPG) manufacturers. Chen and Lee (2012) noted that bullwhip should be measured at the appropriate time unit for cost assessment purposes. However, we conclude that the implications of the aggregation effects may be even more far reaching since it is common for

CPG manufacturers to plan at aggregate levels. In recent years, customer demand planning has become increasingly popular (Lapide 2005), where demand planners create product-family demand plans specific to their highest volume customers, usually on a monthly basis. Our results suggest that customer demand planning processes occurring at such levels of aggregation may be biased due to aggregation effects. By measuring aggregate bullwhip or order variance, planners may be unable to ascertain the true level of variability experienced at operational levels of the organization (Chen & Lee 2012), creating a potential misalignment between future supply and demand.

Additionally, planners may be unable to accurately assess the value of downstream demand signals. That is, aggregate bullwhip measures may indicate that retail customers smooth demand, resulting in order variance that is lower than consumer sales variance. Together, the analytical results of Chen and Lee (2012) along with the empirical results of this research, suggest that planners should make planning decisions at the level with which transactions between the retailer and supplier occur.

Furthermore, planners must consider the effect of seasonality on bullwhip measures. In accordance with Chen and Lee (2012), our results illustrate the dampening effect of seasonality. This is particularly relevant to suppliers seeking to plan based on point of sales data. Since retailers generally attempt to smooth predictable demand fluctuations such as seasonality (Cachon et al. 2007; Bray and Mendelson 2012), order data for seasonal products reflects such smoothing policies and may be more valuable from a planning perspective.

Our results also highlight the substantial difference between the fractional growth rate bullwhip measure and the coefficient of variation and the unit variance bullwhip measures. The fractional growth rate differs from the other two measures in that both the coefficient of variation and unit variance measures are estimated in levels, rather than as percent change. In our data, the percent change in weekly sales is relatively stable while the percent change in orders is usually much larger. In many periods, the percent change in sales is less than one percent, causing instability in the measures. While we note that this issue is likely not relevant to industry- and possibly firm-level analyses, it may cause instability in a measure in the more disaggregate analyses.

H. REFERENCES

Amemiya, T., & Wu., R.Y. (1975). The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association,* 67 (39), 628-632.

Ang, S., Slaughter, S., & Ng., K.Y. (2002). Human capital and institutional determinants of information technology compensation. *Management Science,* 48 (11), 1427-1445.

Blackburn, J. D. (1991). The quick response movement in the apparel industry: A case study in time-compressing supply chains. In: Blackburn, J.D.(Ed.). *Time Based Competition: The Next Battleground in American Manufacturing*. Irwin, Homewood, IL (Chapter 11).

Blanchard, O. J. (1983). The Production and inventory behavior of the American automobile industry. *Journal of Political Economy,* 91 (3), 365-400.

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

Brewer, K. R. W. (1973). Some consequences of temporal aggregation and systematic sampling for ARIMA and ARMAX models. *Journal of Econometrics*, 17 (1), 133-154.

Cachon, G. P., Randall, T., & Schmidt. G.M. (2007). In search of the bullwhip effect. *Manufacturing and Service Operations Management*, 9 (4), 457-479.

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

DeHoratius, N., & Raman, A. (2008). Inventory record inaccuracy: An empirical analysis. *Management Science*, 54 (4), 627-641.

Dejonckheere, J., Disney, S.M., Lambrecht, M.R., & Towill, D.R. (2003). Measuring and avoiding the bullwhip effect: A control theoretic approach. *European Journal of Operational Research*, 147 (3), 567-590.

Dooley, K. J., Yan, T., Mohan, S., & Gopalakrishnan, M. (2009). 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.

El-Beheiry, M., Wong, C.Y., & El-Kharbotly, A. (2004). Empirical quantifaction of bullwhip effect (with application to a supply chain). *Proceedings of the 13th Working Seminar on Production Economics*, *IGLS,* 3: 83-95.

Fransoo, J.C., & Wouters, M.J.F. (2000). Measuring the bullwhip effect in the supply chain. *Supply Chain Management: An International Journal*, 5(2), 78-89.

Gaynor, P.E., & Kirckpatrick, R.C. (1993). *Introduction to Time-Series Modeling and Forecasting in Business and Economics*. McGraw-Hill, Inc, New York, NY.

Ghali, M.A. (1987). Seasonality, aggregation and the testing of the production smoothing hypothesis. *The American Economic Review*, 77 (3), 464-469.

Hammond, J. (1994). Barilla SpA (a). HBS Case 9-694-046, Harvard Business School, Boston, MA.

Homlstrom, J. (1997). Product range management: A case study of supply chain operations in the European grocery industry. *International Journal of Supply Chain Management*. 2 (8), 107-115.

Kahn, J.A. (1987). Inventories and the volatility of production. *The American Economic Review,* 77 (4), 667-679.

Kaipia., R., Korhonen, H., & Hartiala, H. (2006). Planning nervousness in a demand supply network: An empirical study. *International Journal of Logistics Management*, 17 (1), 95-113.

Kleinbaum, D.G., Kupper, L.L., Muller, K.E., & Nizam, A. (1998). *Applied Regression Analysis and Other Multivariable Methods.* Duxbury Press, Pacific Grove, CA.

Lapide, L. (2005). Account level forecasting needs downstream data. *Journal of Business Forecasting*, 24 (2), 17-19.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43 (4), 546-559.

Liao, H., & Chuang, A. (2004). A multilevel investigation of factors influencing employee service performance and customer outcomes. *Academy of Management Journal,* 47 (1), 466-485.

McCullen, P., Towill, D.R. (2002). Diagnosis and reduction of bullwhip in supply chains. *Supply Chain Management: An International Journal*, 7 (3), 164-179.

McKenney, J. L., & Clark, T.H. (1995). Campbell Soup Co.: A leader in continuous replenishment innovations, HBS Case 9-195-124, Harvard Business School, Boston, MA.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical Linear Models.* Sage Publishing, Thousand Oaks, CA.

Rossana, R.J., & Seater, J.J. (1995). Temporal aggregation and economic time series. *Journal of Business & Economic Statistics,* 13 (4), 441-451.

Sobel, M.J. (1969). Production smoothing with stochastic demand I: Finite horizon case. *Management Science*, 16 (3), 195-207.

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

Tiao, G.C. (1972). Asymptotic behavior of temporal aggregates of time series. *Biometrika*, 59 (3), 525-531.

Torres, O.C., & Maltz, A.B. (2008). Understanding the financial consequences of the bullwhip effect in a multi-echelon supply chain. *Journal of Business Logistics*, 31 (1), 23-41.

Wei, W.W.S. (1978). Some consequencoes of temporal aggregation in seasonal time series models. *Seasonal Analysis of Economic Time Series*, Washington D.C., U.S. Department of Commerce, Bureau of the Census.

Table 1 – Product category dimensions

Category	Annual Sales	Seasonality	Life Cycle Stage	Shelf Life
Non-seasonal Dry	$6 billion	None	Mature	Medium
Seasonal Dry	$4 billion	High	Decline	Long
Refrigerated	$2.5 billion	None	Growth	Short

Table 2 – Descriptive statistics

	Fractional Growth Rate	Coefficient of Variation	Unit Variance
Product			
Weekly	331.963	3.117	12.913
Monthly	17.447	1.713	3.780
Category			
Weekly	64.820	2.397	7.250
Monthly	6.357	1.540	3.130

Table 3 – Mean bullwhip comparison

Fractional Growth Rate (e.g., Cachon et al. 2007)

Product-level	Weekly			Four-week		
	Minimum	Mean	Maximum	Minimum	Mean	Maximum
Non-Seasonal	39.86	216.35	725.46	2.37	21.76	143.76
Seasonal	83.04	262.26	501.06	2.95	23.35	115.05
Perishable	8.84	517.28	1205.80	2.38	7.23	22.53
Category-level	**Weekly**			**Four-week**		
	Minimum	Mean	Maximum	Minimum	Mean	Maximum
Non-Seasonal	23.81	41.55	73.96	4.64	10.35	23.44
Seasonal	55.61	100.85	171.18	2.10	4.59	9.32
Perishable	19.96	52.06	125.02	2.52	4.13	7.54

Coefficient of Variation (e.g., Fransoo and Wouters 2000)

Product-level	Weekly			Four-week		
	Minimum	Mean	Maximum	Minimum	Mean	Maximum
Non-Seasonal	1.69	3.85	7.82	0.96	2.29	4.61
Seasonal	1.36	2.23	4.39	0.86	1.29	2.11
Perishable	1.21	3.27	5.63	0.91	1.56	3.75
Category-level	**Weekly**			**Four-week**		
	Minimum	Mean	Maximum	Minimum	Mean	Maximum
Non-Seasonal	1.84	3.29	4.28	1.11	2.18	2.94
Seasonal	1.40	1.78	3.48	1.02	1.27	2.61
Perishable	1.46	2.12	3.30	0.87	1.17	1.44

Unit Variance Ratio (e.g., Torres and Maltz 2008; Chen and Lee 2012)

Product-level	Weekly			Four-week		
	Minimum	Mean	Maximum	Minimum	Mean	Maximum
Non-Seasonal	3.30	18.13	57.04	0.48	5.73	29.47
Seasonal	2.23	5.93	17.00	0.07	1.93	10.90
Perishable	1.59	14.68	46.17	0.47	3.68	14.10
Category-level	**Weekly**			**Four-week**		
	Minimum	Mean	Maximum	Minimum	Mean	Maximum
Non-Seasonal	3.50	11.71	18.48	1.27	5.44	12.34
Seasonal	1.92	2.84	5.52	0.93	1.40	2.08
Perishable	2.53	7.20	17.95	0.83	2.55	10.09

All differences in bullwhip means for product-location and temporal levels are significant at $p<0.01$

Table 4 – Hierarchical linear modeling results for three measures of bullwhip ratios

	Fractional Growth Rate		**Coefficient of Variation**		**Unit Variance**	
Variable	Null	Conditional	Null	Conditional	Null	Conditional
Fixed effects						
Intercept	3.788 ***	2.349 ***	0.754 ***	0.153	1.55 ***	0.601 ***
	0.104	0.395	0.06	0.118	0.134	0.255
Tagg		-2.107 ***		-0.410 ***		-1.087 ***
		0.081		0.017		0.040
Pagg		-1.339 ***		-0.260 **		-0.439 **
		0.264		0.115		0.201
SEAS		-0.258 ***		-0.258 ***		-0.529 ***
		0.065		0.023		0.047
AF		2.086 ***		0.031		-0.223
		0.335		0.087		0.204
Random effects						
DC	-	-	0.018 ***	0.021 ***	0.076 ***	0.109 ***
			0.005	0.004	0.03	0.020
Category	0.194 **	0.148 **	0.083 ***	0.031 ***	0.419 ***	0.087 ***
	0.076	0.065	0.025	0.011	0.127	0.032
Residual	2.221 ***	1.064 ***	0.102 ***	0.044 ***	0.638 ***	0.255 ***
	0.117	0.057	0.006	0.003	0.04	0.017
-2 Log Likelihood	2726.215	2189.763	587.535	65.114	1911.59	1337.950
x^2		536.452 ***		522.421 ***		573.640 ***

Notes: Standard errors are shown in parentheses. All dependent variables and SEAS are transformed into their natural log form. *p<0.1, **p<0.05, ***p<0.01.

Table 5 – Results summary

Influencing factors	Fractional Growth Rate	Coefficient of Variation	Unit Variance
Product-Location Aggregation	Masking	Masking	Masking
Temporal Aggregation	Masking	Masking	Masking
Seasonality	Dampening	Dampening	Dampening

Chapter 4 – TEMPORAL AGGREGATION AND ITS MODERATING EFFECT ON THE VALUE OF POINT-OF-SALE INFORMATION IN FORECAST ACCURACY

IV. TEMPORAL AGGREGATION AND ITS MODERATING EFFECT ON THE VALUE OF POINT-OF-SALE INFORMATION IN FORECAST ACCURACY

A. INTRODUCTION

Accurate demand forecasts are vital to achieving supply chain efficiency and effectiveness given that many operational decisions are based on such forecasts. To reduce forecast error, firms often invest significant resources in sophisticated information systems (Ravichandran & Liu, 2011) designed to generate statistical forecasts (Rexhausen et al., 2012) and facilitate information sharing between supply chain partners (Schoenherr & Swink, 2012). In the retail supply chain, the sharing of consumer sales data (i.e., point-of-sale data) has become increasingly commonplace between retailers and their suppliers. While these firms have invested heavily in systems to improve forecast performance, evidence suggests that the challenge may be only increasing. By 2021, companies will have access to over 35 zettabytes of data generated from supply chain activities (Cognizant, 2012). Of course, utilizing this wealth of data to anticipate future demand is a high priority (Cecere, 2012). Wal-Mart, for example, collects more than 2.5 petabytes of data every hour from its customer transactions (McAfee & Brynjolfsson, 2012), which is in turn shared with suppliers to improve supply chain planning.

Clearly, forecasting with such a high volume of data requires automated processes to generate statistical forecasts; however, for these processes to function effectively, managers must properly design them in light of existing theory. One critical decision is the determination of the level of temporal aggregation at which statistical forecasts should be generated. Temporal aggregation is the process where a high frequency time series (e.g., weekly) is aggregated to a lower frequency time series (e.g., monthly) (Nikolopoulos et al., 2011). Many manufacturers in the retail supply chain face the conundrum of whether to forecast retailer requirements in weekly or monthly increments. The conundrum results from the fact that corporate planning processes, such as sales and operations planning (S&OP), recur monthly (Olivia & Watson 2011), yet operational decisions, such as scheduling outbound logistics activities are often made on a weekly basis.

In this study, we draw on analytical models based on statistical theories to inform the consequences of using temporally aggregated data to generate demand forecasts in the context of a retail supply chain. We find that countervailing effects of temporal aggregation may exist, namely information loss and variance reduction. Under statistical theory of information loss (Amemiya & Wu, 1972; Marcellino, 1999), time series properties that inform the underlying data generating process of the time series become altered and lost during the temporal aggregation process. On the other hand, following the premise underlying risk pooling, where random errors are canceled via aggregation, forecasters often assume that temporal aggregation will likely provide more stable and accurate forecasts and thus prefer to create statistical forecasts using temporally aggregated data (Finn, 2004; Hotta et al., 2005).

To reconcile these competing notions, both based on established statistical concepts, we hypothesize that the dominant effect of temporal aggregation likely depends on the information source being used by the supplier to create the demand forecasts. That is, the decision of whether to temporally aggregate data should be dependent on if the supplier uses shared retail sales information or historical order information to forecast future retailer requirements. We suggest that this is due to the relative levels of randomness and the degree of autocorrelation inherent in these information types. To test our hypotheses, we design a quasi-experiment based on data from two highly shopped grocery categories where both the level of temporal aggregation and the information type utilized are manipulated. To analyze the quasi-experimental results, we utilize a hierarchical linear model (HLM) due the multi-level nature of the data.

Moving forward, we review the forecasting and information sharing literature to set the context for our study. Next, we develop our hypotheses using both the supply chain and econometrics literatures and analytically show how temporal aggregation may transform statistical properties of a time series but also reduce its variance. We then utilize weekly observations of paired order and POS data over a period of two years to design a quasi-experiment to generate competing forecast for conditions of temporal aggregation and information type. Section 4 outlines our empirical method to test our hypotheses. Following the description of our data and analysis, we present our results, draw conclusions from the study, and offer managerial implications.

B. THEORY AND HYPOTHESES DEVELOPMENT

<u>Retail Forecasting and Replenishment</u>

In the retail supply chain, retailers amass an incredible amount of data captured through customer transactions. These data serve as the basis for vital decisions associated with inventory, storage, and replenishment (Schmarzo, 2012) through timely demand forecasts and planning (McAfee & Brynjolfsson, 2012). Over time, advancement in supply chain management strategies gave rise to increasingly collaborative demand forecast and planning practices such as S&OP that place particular emphasis on information sharing to drive operational efficiencies (e.g., Cachon & Fisher, 2000).

The replenishment process to retail stores is typically accomplished through either direct-to-store delivery (DSD) or through the retailer's network of distribution centers (DCs). If retail DCs are utilized, the process generally follows a model where a set of individual retail stores place orders to a single regional DC. The DC, in turn, places orders on a periodic basis to a supplier DC. The supplier is then responsible for replenishing a specified set of retail DCs. For suppliers, the orders placed by the retail DCs, (DC orders), are of particular interest as they represent its customer (not consumer) demand. DC orders, as compared to POS, are difficult to forecast accurately. This is due to the fact that the variance of DC orders is most often greater than the variance aggregate sales recorded by the stores replenished by the DC (i.e., the bullwhip effect).

It is easy to assume that a DC's orders might be easily predicted by summing the sales of the stores replenished by the particular DC. Yet, "store replenishment and execution processes, retailer distribution center (DC) replenishment and operating procedures (Vogt, 2010; Kum, Balakrishnan, & Chun, 2010), warehouse management system idiosyncrasies (Autry et al., 2005), and other supply chain processes, such as postponement (Zinn & Bowersox, 1988), inventory centralization (Evers, 1995; 1996; 1997; Evers & Beier, 1993; 1998; Mahmoud, 1992; Ronen, 1990; Tallon 1993; Zinn et al. 1989) and lean practices (Goldsby, Griffis & Roath, 2006) introduce complexity into the retailer's ordering processes" (Williams & Waller 2010, pg. 1), making DC orders more variable and difficult to predict than retail sales.

To forecast DC orders, consumer packaged goods (CPG) suppliers generally use a simplistic process. Time series forecasting methods, like exponential smoothing forecast future customer demand based on the archived order data (Williams & Waller 2010). Very often, customer demand forecasts predict each retail DC's requirements in weekly intervals for each DC in order to make operational decisions. In order to quickly respond to short-term market conditions, retailers typically order from suppliers on a weekly basis (Nijs et al. 2007).

Therefore, suppliers often base transportation and labor capacity decisions on these weekly customer demand forecasts (e.g., Cachon & Fisher 1997).

Recent developments in point-of-sale (POS) sharing have increased interest in whether POS might improve the ability of the supplier to forecast customer demand. Williams and Waller (2010) compare the forecast accuracy of customer demand forecasts based on POS history with those based on order history and find that POS-based forecasts outperform those based on order history in approximately 65% of the cases; however, order history outperforms POS in the remaining 35%, indicating POS and order history may have unique information that can help predict future customer ordering behavior.

Temporal Aggregation and Information Loss

The effect of temporal aggregation on time series has been studied for decades, beginning with the seminal work of Amemiya and Wu (1972), which investigates the issue of information loss. Amemiya and Wu study temporal aggregation where the data is an autoregressive (AR) process of order *p*. The literature later generalizes the effect of temporal aggregation to include autoregressive moving average models with exogenous variables (ARMAX) (Brewer 1973), seasonal structures (Wei 1978) and nonstationary models (Tiao 1972). For a complete overview of temporal aggregation techniques, we refer the reader to Silvestrini and Veredas (2008).

Information loss refers to a loss of information about the underlying data generating process of the time series. To illustrate how temporal aggregation may result in such information loss, we model the effect of temporal aggregation where a time series (x_t) is a first-order autoregressive process, $x_t \sim AR(1)$, which can be expressed as,

$$x_t = \varphi_1 x_{t-1} + \varepsilon_{1,t}, \text{ where } \varepsilon_{1,t} \sim (0, \sigma_1^2) \tag{1}$$

The expected value and variance of the above expression are known to be:

$$E[x_t] = \frac{\mu_1}{1-\varphi_1}, \text{ and } Var[x_t] = \frac{\sigma_1^2}{1-\varphi_1^2} \quad (2)$$

Temporal aggregation occurs when demand observed at two consecutive, non-overlapping periods are summed. Therefore, the summation of x_t and x_{t+1} may be expressed in an aggregated time series z_t, where $z_T \sim ARMA(1,1)$and can be defined such that[3]:

$$z_T = \varphi_2 z_{T-1} + \theta\varepsilon_{2,T}, \text{ where } \varepsilon_{2,T} \sim (0, \sigma_2^2) \quad (3)$$

Therefore,

$$E[z_T] = \frac{\mu_2}{1-\varphi_2}, \text{ and } Var[z_T] = \frac{\sigma_2^2(1+\theta^2-2\varphi_2\theta)}{1-\varphi_2^2}. \quad (4)$$

From this simple model[4], we find that the aggregated time series (z_t) has a different underlying data generating process than the disaggregate time series (x_t).

In a retail supply chain context, aggregation of demand signal data from weekly to monthly eliminates information such as paycheck cycles. For example, many firms in the U.S., including the government, pay their employees on a bi-weekly basis. As a result, retail sales tend to follow a similar pattern. Thus, weekly patterns in weekly customer requirements can be masked as weekly demand signal data gets aggregated into monthly data. Therefore, practical evidence suggests and statistical theory predicts that the temporal aggregation of a time series results in information loss about the underlying data process which may have severe negative implications for prediction of future observations of the time series (Rosanna & Seater 1995).

Temporal Aggregation and Variance Reduction

While it seems that temporally aggregated time series cannot be better predictors than disaggregate predictors (Amemiya & Wu 1972), practitioners often tend to prefer using

---

[3] We refer readers to Brewer (1973) for its derivation.
[4] According to Tiao (1972), φ_1, φ_2, and θ are independent of each other.

temporally aggregated data to forecast at disaggregated levels for reasons such as simplicity (Finn, 2004). Despite the mounting analytical evidence that temporal aggregation results in substantial loss in information in high frequency data (e.g., Amemiya & Wu, 1972; Brewer, 1973; Wei, 1978; Nijman and Palm, 1990), we note that statistical theory "may not be definitive because some of the results are asymptotic and leave open the question of what happens with actual data" (Rossana & Seater 1995, p. 443).

In fact, we argue that a countervailing effect to the information loss effect of temporal aggregation exists, namely, variance reduction. A major contributing factor to forecast error throughout the retail supply chain is the bullwhip effect. The bullwhip effect is defined as the amplification of order variance as orders move from the retail echelon to the manufacturing echelon of the supply chain (Lee et al., 1997). As retail sales translate into orders placed by stores to the supplying distribution center (DC), and then on to a supplier's DC, the variance of orders is amplified at each echelon, resulting in a more "noisy" demand signal. Often, the causes of the increased variability are not related to demand factors but managerial and behavioral idiosyncrasies that are not useful information for forecasters when predicting future retailer inventory requirements (Lee et al. 1997). For example, a retailer that stockpiles inventory will likely place future order of zero as the stockpiled inventory sells down.

To deal with the increased variability, forecasters often temporally aggregate demand signal data to reduce the data's variance (Finn, 2004; Nikolopoulos et al., 2011). This practice is based in the statistical concept of risk pooling which underlies the portfolio effect in the supply chain management literature. Similar to variance reduction achieved by consolidation of inventory holding locations (Zinn et al.1989; Ronen 1990; Mahmoud 1992; Evers & Beier 1993,1998; Tallon 1993; Evers 1995, 1996, 1997; Das & Tyagi 1999; Ballou 2005) and product aggregation (Williams & Waller 2011), the temporal aggregation of demand signal data partially cancels the random errors in the time series. To model the variance reduction due to temporal aggregation, we again consider the effect where a time series (x_t) is a first-order autoregressive process,$x_t \sim AR(1)$, which can be expressed as,

$$x_t = \varphi x_{t-1} + \varepsilon_t, \text{ where } \varepsilon_t \sim (0, \sigma_t^2) \tag{5}$$

Note that σ_t represents the standard deviation of the time series at time t while σ_{t-1} represents the standard deviation of the time series at $t-1$. We assume that errors are homoscedastic, thus $\sigma_t = \sigma_{t-1}$.

We define $\{X_{t+(t-1)}\}$ as a non-overlapping aggregated demand series, where $X_{t+(t-1)} = X_t + X_{t-1}$. To examine whether variability is reduced through temporal aggregation, i.e., whether $\sigma_{t+(t-1)} < 2\sigma_t$, we express the standard deviation of $X_{t+(t-1)}$ as,

$$\sigma_{t+(t-1)} = \sqrt{\sigma_t^2 + \sigma_{t-1}^2 + 2\theta\sigma_t\sigma_{t-1}} \tag{6}$$

Since $\sigma_t = \sigma_{t-1}t$ due to constant variance, equation 6 can be rewritten as,

$$\sigma_{t+(t-1)} = \sqrt{\sigma_t^2 + \sigma_t^2 + 2\theta\sigma_t\sigma_t} \tag{7}$$

Equation 7 can be further simplified algebraically, expressed as,

$$\sigma_{t+(t-1)} = \sigma_t\sqrt{2(1+\theta)} \tag{8}$$

From equation 8, we can observe that if $\theta = 1$ (i.e., perfect positive autocorrelation), then $\sigma_{t+(t-1)} = 2\sigma_t$. Otherwise, for all $\theta < 1$, $\sigma_{t+(t-1)} < 2\sigma_t$. In addition, we may also observe the following properties for equation 8:

$$\lim_{\theta \to 1} \sigma_{t+(t-1)} \to 2\sigma_t$$

$$\text{Otherwise, } \sigma_{t+(t-1)} < 2\sigma_t,\ for\ all -1 \le \theta < 1 \tag{9}$$

That is, we observe that temporal aggregation results in reduced variance as long as the time series is not perfectly, positively autocorrelated, and the degree of variance reduction is dependent upon the degree of autocorrelation in the time series.

Temporal Aggregation and Information Source

While aggregation is a method to deal with the amplified variance in upstream demand signals, another method is to utilize a downstream demand signal to forecast customer requirements. Information sharing is an important enabler of collaboration in the retail supply chain (Barratt & Barratt, 2011; Schoenherr & Swink, 2012). Recent advances in information technology increased the sharing of retail POS to provide to suppliers the option of forecasting customer demand, using either POS or order history. As shown by Williams and Waller (2010), POS is generally the preferred data, because of the lower levels of variance, relative to order data.

Since a key distinction between POS and DC order data is the associated variance, we argue that the effect of temporal aggregation on these demand signals may differ given countervailing effects of temporal aggregation. Given that order data tend to have high levels of variance due to the bullwhip effect, we expect that the variance reduction due to the pooling of observations may be the dominant statistical effect when temporally aggregating order data and, as the random errors are canceled,, forecast error is likely to decrease.

On the contrary, POS data is not subject to the bullwhip effect and tends to have lower variance than order data. Therefore, the potential benefit of variance reduction due to temporal aggregation is much less. In fact, we contend that the information loss effect may be dominant when temporally aggregating POS data. That is, as POS data is temporally aggregated, the loss of information about the underlying nature of consumer sales has a negative effect on the ability to forecast customer demand and overshadows any potential benefit of variance reduction.

Given the countervailing statistical effects of temporal aggregation on these demand signals, we hypothesize that the effect of temporal aggregation on order forecast error differs upon whether the supplier utilizes order or POS data to forecast customer demand. Therefore, we hypothesize an interaction effect between temporal aggregation of the demand signal data and whether POS or order data is used to generate the statistical forecast, such that:

> H_{1a}: Temporal aggregation is positively associated with customer demand forecast error, when POS data is utilized to generate statistical forecasts.
>
> H_{1b}: Temporal aggregation is negatively associated with customer demand forecast error, when order data is utilized to generate statistical forecasts.

Autocorrelation and Temporal Aggregation

While we anticipate that temporal aggregation compromises the benefit of using POS data to generate customer demand forecast (as indicated in H_{1a}), this effect can be potentially amplified by the POS data's autocorrelation factor. Since the countervailing effects of information loss and variance reduction are contemporaneous, temporal aggregation is likely to have both effects on POS data. Specifically, as shown in equation 8, the autocorrelation factor plays an important role in governing the statistical effect of variance reduction: As the autocorrelation factor of POS data at the disaggregate level (i.e., weekly) approaches perfect autocorrelation (i.e., $\theta \rightarrow 1$), the variance reduction effect will be minimized at the temporally aggregated level (i.e., monthly).

In addition, the autocorrelation factor's role in variance reduction will also impact temporal aggregation's effect on forecast error utilizing DC order data. In this instance, while we anticipate that variance reduction is a principle benefit of temporally aggregating DC order data, thereby reducing forecast error (as indicated in H_{1b}), this effect is further moderated by the DC order data's autocorrelation factor. Similar to its effect on POS data, as the autocorrelation factor of DC order data at the disaggregate level approaches perfect autocorrelation, the variance reduction effect in this instance will also be minimized

Thus, the autocorrelation factor of both POS and DC order data may influence the impact on forecast error due to the interaction effect of temporal aggregation and information source. As shown above, temporal aggregation results in both altering the statistical properties of POS and DC order data to cause information loss (e.g., Amemiya & Wu, 1972; equation 4) as well as variance reduction (equation 8). For both POS and DC order data, diminishing variance reduction effect due to increasingly positive autocorrelation factor will result in higher forecast errors. Alternatively, for POS and DC order data that have increasingly negative autocorrelation, the heightened variance reduction effect will result in lower forecast errors.

To summarize, autocorrelation factor determines the magnitude of the variance reduction effect. For both POS and DC order data, variance reduced through temporal aggregation is moderated by each data series' autocorrelation factor at the disaggregate level. In both instances, an increasingly positive autocorrelation factor results in diminishing variance reduction effect, thereby increasing forecast error. Formally stated:

> H_{2a}: Autocorrelation factor is positively associated with customer demand forecast error, when POS data is temporally aggregated prior to generating statistical forecast.
>
> H_{2b}: Autocorrelation factor is positively associated with customer demand forecast error, when DC order data is temporally aggregated prior to generating statistical forecast.

C. METHODS AND MEASURES

Quasi-Experimental Design

To test our hypotheses, we design a quasi-experiment for forecast accuracy based on two years of weekly DC order and POS data obtained from a large consumer packaged goods supplier. The forecast experiment compares weekly order forecasts using DC order and POS data at weekly and monthly levels of temporal aggregation.

For short and mid-horizon forecasts of fast-moving items, exponential smoothing techniques are the most commonly utilized in industry (Mentzer & Kahn 1995; McCarthy et al. 2006) and generally offer forecast accuracy competitive against other approaches that are substantially more complex (Makridakis et al., 1982; Makridakis & Hibon 2000). Considering that trend is likely present in the data, for which exponential smoothing alone is not sufficient, we utilize Holt's exponential smoothing with trend.

To setup the quasi-experiment, we first aggregate an initial 88 weekly observations of DC order and POS data to 22 monthly observations. Next, we estimate for each SKU-DC combination the smoothed components for the level and trend for order and POS data at both weekly and monthly levels of temporal aggregation. For our forecasting parameters, we chose three values for α and β (α=0.51; α=0.19; α=0.02; β=0.176; β=0.053; β=0.005) based upon the range of reasonable values (Silver, Pyke and Peterson 1998, p. 108). In addition, not all combinations of forecasting parameters were used. For stability purposes, the value of β should be well below that of α (McClain & Thomas 1973). Thus, we utilize a total of six combinations out of a possible total of nine (0.51, 0.176; 0.51, 0.053; 0.51, 0.005; 0.19, 0.053; 0.19, 0.005; 0.02, 0.005). Further, an initialization of the forecast for the first period is required for the single exponential smoothing method. For the OF-competition, the initial forecast was set to the value of the actual order for the first period (Hanke & Wichern 2005, p. 118).

We next utilize the estimated level and trend components to generate customer demand forecasts for each SKU-DC combination over a 13-week out-of-sample forecast horizon (i.e., fiscal quarter). The calculation for weekly forecast error is straight-forward for order forecast generated using weekly order and POS data. Since order forecast generated using monthly data contains expected orders over four weeks, we divide monthly order forecasts by four to obtain their weekly-equivalent values before proceeding to calculate the weekly forecast error. Thus, our quasi-experiment is in a two-by-two design (Figure 1) for two sources of information for forecast input (DC orders and POS) and two levels of temporal aggregation (weekly and monthly).

Data Collection and Measures

Our data includes DC order and POS data for two grocery categories. The first category is a mature, dry grocery product category and is one of the highest volume grocery categories in a typical supermarket retail format. The second category features fresh, refrigerated products that have short shelf-lives and thus flow through the distribution network relatively quickly.

Our sample includes weekly data for nine dry grocery SKUs and five refrigerated SKUs. The weekly data were collected over a period of two years at six regional U.S. DCs owned by one of the manufacturer's largest retail customers, for a total of 82 unique SKU-DC combinations. DC orders are defined as the weekly orders placed by a particular retail DC to the manufacturer while POS is the cumulative weekly sales of the retail stores replenished by the particular DC.

To evaluate the out-of-sample forecast performance, we measure customer demand forecast error with mean absolute deviation (MAD). MAD measures forecast error by averaging the absolute value of the DC order forecast errors, which is calculated as the difference between actual weekly orders (A) and weekly order forecast ($\widehat{W}$), and is a measure of the magnitude of forecast error. The calculation of MAD_{ijkfm} using weekly level of aggregation is shown:

$$MAD_{ijkfm} = \frac{\sum_{m=1}^{n} \left| A_{ijkfm} - \widehat{W}_{ijkfm} \right|}{n},$$

where n is the number of periods over which the MAD is calculated, I denotes the 14 products across j categories stocked in k DCs, and m indicates the 13 weeks of forecast horizon. In addition, the calculation of MAD_{ijkfm} using monthly order forecast ($\widehat{M}$) is shown:

$$MAD_{ijkfm} = \frac{\sum_{m=1}^{n}\left|A_{ijkfm} - \left(\frac{\widehat{M}}{4}\right)_{ijkfm}\right|}{n}$$

In our quasi-experiment, the variables of interest are binary variables *AGG* and *POS*, where *AGG* = 1 if demand is aggregated at the monthly level and 0 otherwise, and *POS* = 1 if the forecast input is POS data and 0 if order data.

<u>Controls</u>

We code each DC in accordance to its location as *DC*. Our aim is to control for any unmeasured differences among DCs due to managerial or regional idiosyncrasies that might exist. In addition, our products come from two categories each with its unique demand characteristics such as shelf life. Therefore we also code these categories accordingly as variable *CAT*. As previously mentioned, we generated customer demand forecasts using six pairs of reasonable smoothing parameters. Thus, we code a third control variable, *FP*, for each combination of the smoothing parameters (see footnote 2). All of these variables are to be used to control for the potential lack of independence within our dependent variable.

In addition, we include two additional variables that may potentially affect forecast error. First, forecast error tends to be affected by average demand (Mentzer & Cox 1984). Therefore, we include the average weekly demand volume for each product, coded as *Mean*. In addition, autocorrelation can also affect forecast accuracy. Therefore we derive the autocorrelation factor for each product as well, coded as *AR*.

D. ESTIMATION AND RESULTS

Average MAD and Autocorrelation Factor

We present in Figure 2 an evaluation of average MAD broken down by temporal aggregation as well as data source. Initial evaluation of average MAD supports the notion that POS is generally superior to DC orders in forecast accuracy (Williams & Waller, 2010) due to benefits of information sharing (Cachon & Fisher, 2000). However, the difference between DC order-based and POS-based order forecast diminishes as level of temporal aggregation moves from weekly to monthly. Using weekly data, POS-based order forecast demonstrates statistically significant improvement in order forecast error ($F = 38.47$, $p<0.01$) over DC order-based forecast. However, when using monthly data, while POS-based order forecast is still nominally lower than DC order-based forecast, the difference is no longer statistically significant ($F = 0.028$, $p>0.10$).

In comparing mean MAD difference between levels of temporal aggregation, results initially suggest that while the improvement to DC order-based forecast is not statistically significant ($F = 0.767$, $p>0.10$), the increase in POS-based forecast is indeed significant ($F = 3.84$, $p<0.05$). Thus, evaluation of mean forecast error yielded by different cells of our quasi-experiment appears to provide initial support for our hypotheses. Lastly, we note an interesting observation in the average autocorrelation factor for DC orders and POS. While POS exhibits positive autocorrelation ($AR_{POS} = 0.73$), DC orders are instead negatively autocorrelated ($AR_{Order} = -0.18$).

Hierarchical Linear Modeling

Since our data for the 14 products are nested in two categories and six distribution centers, the assumption of independence as required for ordinary least squares (OLS) estimation is violated. Further, our forecast uses six different combinations of forecast parameters based on three levels of alpha and beta values, which results in systematic influence on the calculation of forecast errors. Therefore the traditional ANOVA-based analysis is not appropriate. In order to account for the unobservable, idiosyncratic effects on forecast error, we use hierarchical linear modeling (HLM) to model the multi-level structure of the data (Raudenbush & Bryk 2002). Specifically, HLM parcels out variance components based on higher levels of groups that may exert influence on measurement of the dependent variable.

Hypothesis Testing

We test our hypotheses regarding the interaction of effect of temporal aggregation and the use of POS and order data to forecast customer demand in multiple steps. Similar to the approach of Ang et al. (2002), DeHoratius and Raman (2008), and Liao and Chuang (2004), we estimate our model using full maximum likelihood in three stages. First, we estimate a null model where no control or predictor variables are included. In Model 1, we add the previously described control variables, and in Model 2, the experimental factors and their interaction are included. The HLM results are presented in Table 2.

Null Model

To adequately account for all three potential influences, we begin our empirical investigation by partitioning the dependent variable into the variance across products I (I = 1, …, 14), categories j (j=1,2), DCs k (k=1, … 6) and combinations of forecast parameter f (f=1, … 6). We estimate the model using full maximum likelihood similar to Ang et al. (2002), DeHoratius and Raman (2008), and Liao and Chuang (2004). Thus our null model is:

$$MAD_{ijkf} = \theta_0 + CAT_{000j} + DC_{000k} + FP_{000f} + e_{ijkf}$$

where θ_0 is the fixed intercept parameter, while the random effect parameter of category j is CAT_{000j}, the random effect parameter of DC k is DC_{000k}, and the random effect parameter of combinations f is FP_{000f}. Finally, the random effect parameter of product I is e_{ijkf}. Note that

CAT_{000j} , DC_{000k}, FP_{000f} and e_{ijkf} are each normally distributed with a zero mean and variances of τ_{CAT000}, τ_{DC000}, τ_{FP000}, and σ^2, respectively.

In this estimation, results indicate statistical significance for all three potential sources of structural influence. Category's covariance parameter indicates that it accounted for 72% of the overall variance in *MAD* (τ_{CAT000} = 13,607.57, p < .05). DC effects accounted for an additional 19% of the variance (τ_{DC000} = 3,583.58, p < .01). Forecast parameters accounted for a small but statistically significant 1% of the total variance (τ_{FP000} = 166.61, p < .01). The remaining 8% of the variance in *MAD* may thus be attributed to product level effects.

Conditional Models

For our conditional models, we add the fixed effects for our independent variables to our model hierarchically by entering the control fixed effects first, then our variables of interest and their interaction effect. In Model 1, we enter our control variables to the null model. They include $MEAN_{ijkf}$, AR_{ijkf}, Agg_{ijkf}, and POS_{ijkf}. We then enter to Model 2 our two-way interaction of interest to the model, $Agg_{ijkf}\, x\, POS_{ijkf}$. With the inclusion of this interaction term, we may obtain the estimated marginal means to test effect of temporal aggregation on order forecast error when DC order data is used (H_{1b}) as well as when POS data is used (H_{1a}). Finally, we enter in Model 3 the full factorial of two-way interaction effects for AR_{ijkf}, Agg_{ijkf}, and POS_{ijkf} as well as their three-way interaction in order to test the moderating influence of autocorrelation factor on the interaction effect of temporal aggregation and the use of POS data. Thus, our full conditional model is specified as,

$$MAD_{ijkf} = \theta_0 + CAT_{000j} + DC_{000k} + FP_{000f} + e_{ijkf}$$

$$+\beta_1 MEAN_{ijkf} + \beta_2 AR_{ijkf} + \beta_3 AGG_{ijkf} + \beta_4 POS_{ijkf}$$

$$+\beta_5 AGG_{ijkf} * POS_{ijkf} + \beta_6 POS_{ijkf} * AR_{ijkf} + \beta_7 AGG_{ijkf} * AR_{ijkf}$$

$+\beta_8 AGG_{ijkf} * POS_{ijkf} * AR_{ijkf}$

where β_1 and β_2 are the fixed effects on MAD_{ijkf} due to the mean and underlying autocorrelation factor of POS demand, respectively. B_3 and β_4 are the fixed effects across products for temporal aggregation and the use of POS history as forecast input, respectively. B_5 to β_7 are the fixed effects across products for the two-way interaction effects among AR_{ijkf}, Agg_{ijkf}, and POS_{ijkf}. Finally, β_8 is the estimated three-way interaction effect.

Results

As expected, an analysis reveals a significant interaction ($\beta_5 = 16.66, p < 0.000$; see Table 2, Model 2) between the demand signal used to forecast and temporal aggregation from weekly to monthly observations. The simple effect analyses from the HLM results indicate that aggregation from weekly to monthly when using POS data to generate statistical forecasts significantly increases MAD from to 135.548 to 144.685 ($F = 12.830$, $p < 0.000$), providing support for H_{1a}. On the contrary, the HLM results indicate that aggregation from weekly to monthly when using order data to forecast significantly decreases MAD from 150.872 to 143.349 ($F = 18.640$, $p < 0.000$), providing support for H_{1b}.

H_{2a} concerned the three-way interaction between temporal aggregation, the use of POS data, and autocorrelation factor. Model 3 shows the estimated coefficients β_7 and β_8, which is the estimated effect of autocorrelation factor on forecast error when temporally aggregated DC order data is utilized ($AGG_{ijkf} x\ AR_{ijkf}$) and when temporally aggregated POS data is utilized ($AGG_{ijkf} x POS_{ijkf} x AR_{ijkf}$), respectively. We find that AR_{ijkf}significantly increases forecast error in both instances: When temporally aggregated DC order data is used, forecast error increases if autocorrelation is positive ($\beta_7 = 19.24, p < 0.000$); Similarly, when temporally aggregated POS data is used, forecast error also increases if autocorrelation is ($\beta_8 = 35.56, p < 0.000$). Thus, Model 3 offers support for both H_{2a} and H_{2b}.

E. DISCUSSIONS OF MANAGERIAL AND THEORETICAL IMPLICATIONS

Our findings confirm the existence of countervailing statistical effects due to temporal aggregation in the context of customer demand forecasting using different demand signals. By plotting the interaction effect (Figure 3) between temporal aggregation and the information type used to generate the customer demand forecasts, we gain further insight into this issue. The plot clearly reveals that utilizing temporally aggregated data to forecast at the disaggregated level (i.e., a temporally top-down approach) has opposing effects on customer demand forecast error depending upon which information type is utilized. That is, a temporally top-down approach to forecasting increases forecast error when POS data is utilized, but decreases forecast error when DC order data is utilized.

In addition, we further find that the above temporal aggregation-information type interaction effect is moderated by the underlying autocorrelation factor of the information, as illustrated by Figures 4 and 5. For both POS and DC orders, as the autocorrelation factor becomes increasingly negative, the temporally top-down forecasting approach decreases forecast error as the variance reduction effect is increased. However, as the autocorrelation factor becomes increasingly positive, the temporally top-down forecasting approach increases forecast error as adverse effects from information loss overtakes the benefit of variance reduction.

Finally, Figure 6 illustrates the full picture of the three-way interaction, which reveals two significant findings. First, the superior information content embedded in POS makes it the preferred demand signal for forecasting. However, while monthly POS is the preferred forecast input at low levels of autocorrelation, weekly POS instead offers superior forecast accuracy when autocorrelation is strongly positive. Second, variance reduction effect of temporal aggregation reduces forecast error regardless of information source.

Managerial Implications

Suppliers require accurate and timely forecasting to properly position inventory throughout its distribution network and schedule outbound logistics transportation operations. Often, suppliers may choose to temporally aggregate input data under the belief that it improves forecast accuracy. Our results indicate that temporal aggregation has two countervailing effects, namely variance reduction and information loss. Whereas the variance reduction effect reduces forecast error, information loss increases it. Their collective effect on forecast error then depends on the autocorrelation factor of the forecast input. As autocorrelation factor becomes increasingly positive, information loss effect overtakes variance reduction effect to increase forecast error. As a result, temporal aggregation can either improve or harm forecast accuracy. Our findings have a clear managerial implication that is relevant to most suppliers in the retail supply chain.

Recent industry consolidation had left only a small number of customers who account for increasingly large portions of suppliers' total volumes (Hofer et al., 2012). As a result, suppliers are becoming increasingly reliant on utilizing key customer account forecasts (Lapide 2007). The most readily observed customer demand signal from the suppliers' perspective is the order data of their customers (i.e., the retailers), and as a result contains potentially valuable information indicating customer order behavior (Williams and Waller, 2010). But these orders are prone to high degrees of fluctuation as due to the bullwhip effect (Lee et al. 1997). Our results suggest that suppliers may counter this phenomenon through selective use of temporally top-down forecasting approach. If a customer's demand signal is negatively autocorrelated, greater variance reduction effect will result in superior demand forecast. Conversely, variance reduction effect will diminish to yield less accurate forecast.

On the other hand, increasing number of retailers is sharing their POS data with suppliers. As a result, suppliers have visibility to demand signals from both their retail customers as well as consumers. Our results once again show that POS data can be a superior source of information. Specifically, POS data tends to have less variance and provide suppliers with a more accurate view of consumer demand (Williams & Waller, 2010). However, effective utilization of POS for forecasting is dependent on the tactical use of temporally top-down forecasting approach. If POS is negatively autocorrelated, then suppliers should aggregate their data prior to forecasting to take advantage of the substantial benefit associated with variance reduction effect despite potential loss of information. If POS is positively correlated, then suppliers should avoid temporal aggregation, since doing so will result in substantial loss in information that is not outweighed by the benefit from variance reduction. Lastly, suppliers should be particularly cautious with the use of negatively autocorrelated POS data. Our results indicate that if such data is not temporally aggregated, forecast errors can be higher than those obtained with DC order data.

<u>Theoretical Implications</u>

A long line of analytical literature on temporal aggregation argues that aggregation results in information loss to lead to decreased forecast accuracy (Amemiya & Wu, 1972; Rossana & Seater, 1995). Yet empirical studies frequently concluded to the contrary (e.g., Hotta et al., 2005). We contribute to this discussion by showing that both variance reduction and information loss exist simultaneously in the temporal aggregation process. While temporal aggregation can improve forecast accuracy through variance reduction, this effect is dependent on the autocorrelation factor of the data series—as data becomes increasingly positively correlated over time, the effect of variance reduction diminishes. On the other hand, while information loss also occurs, if data is sufficiently negatively autocorrelated, the benefit of variance reduction effect can outweigh information loss to improve forecast accuracy.

In the context of the greater information sharing and supply chain management literature, this study has broader implications. Data is being generated at increasingly higher volumes (McAfee & Brynjolfsson, 2012). As a result, firms invest significant resources in sophisticated inter-organizational information systems (Ravichandran & Liu, 2011) to automate the collection, utilization, and dissemination process. This study demonstrates that a key to the efficient and effective use of data is the proper specification of temporal aggregation prior to the forecasting process. In particular, a chief benefit of information sharing is to enable firms to collaborate with supply chain partners (Allred et al., 2011) by adopting one-number forecasting (Finn, 2004) to synchronize supply chain activities (Cao & Zhang, 2011). Whereas the extant literature emphasizes how data is collected and shared, our results indicate that how such information is technically processed can have impact on their utility.

One of the keys to the efficient and effective use of shared data is to carefully consider the interaction between temporal aggregation and information source: While temporal aggregation can lower variance to improve forecast accuracy, it can also mask valuable information such as consumer demand patterns. In the collaboration process, operations planning idiosyncrasies among firms and functions frequently result in conflicting levels of temporal aggregation at which data is collected and utilized (Pauwels et al., 2004). Thus, as firms engage in collaborative demand planning to generate statistical forecasts (Rexhausen et al., 2012), the proper selection of forecast parameters such as level of temporal aggregation and information source can result in improved forecast accuracy.

F. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

In this study, we reconciled two conflicting effects of temporal aggregation on forecast accuracy by utilizing paired order and POS data collected for a large number of SKU-DC combinations. Whereas analytical literature generally argue that temporal aggregation results in less accurate forecast, a large number of empirical studies and evidence from industry point to the contrary. We find that temporal aggregation's effect on forecast accuracy in the retail supply chain is dependent on the source of the input data. However, we note some limiting factors that should be pursued. First, our data come from two high volume, non-seasonal categories. Further research should address our research questions in both seasonal as well as low-volume categories. In addition, since our data come from one single retailer, we are unable to assess the potential differences between retail formats as well as pricing strategies.

Additional research could also examine additional types of forecast methods. While we used the most commonly-utilized time series forecast methods given our category characteristics, other more complex (and simpler) quantitative and qualitative forecast methods can yield additional insight in collaborative demand planning in the supply chain. For example, bullwhip can result from both deliberate as well as random managerial and behavioral idiosyncrasies (Lee et al., 1997). Future research can attempt to parcel out the incremental variance due to deliberate managerial policies that are predictable (e.g., planned inventory build-up) from those that are random (e.g., gaming for fear of shortage). Alternatively, category growth implies a moving average. Hence, an increasing mean with corresponding increase in variance may give companies additional incentive to use the multiplicative model. Otherwise a data series with predominant growth in mean without matching increase in variance may instead yield lower forecast error through the use of the additive model.

G. REFERENCES

Allred, C.R., Fawcett, S.E., Wallin, C., & Magnan, G.M. (2012). A dynamic collaboration capability as a source of competitive advantage. *Decision Sciences,* 42, 129-161.

Amemiya, T., & Wu., R.Y. (1975). The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association,* 67 (39), 628-632.

Anand, S., & Cunnane, C. (2009). *Inventory Optimization: Retail Strategies for Eliminating Inventory Stock-Outs and Over-Stocks,* Aberdeen Group.

Ang, S., Slaughter, S., & Ng, K.Y. (2002). Human capital and institutional determinants of information technology compensation. *Management Science,* 48, 1427-1445.

Ballou, R.H., (2005). Expressing inventory control policy in the turnover curve. *Journal of Business Logistics*, 26, 143-164.

Barratt, M., & Barratt, R. (2011). Exploring internal and external supply chain linkages: evidence from the field. *Journal of Operations Management,* 29, 514-528.

Brewer, K.R.W. (1973). Some consequences of temporal aggregation and systematic sampling for ARIMA and ARMAX models. *Journal of Econometrics*, 1, 133-154.

Brannas, K., & Ohlsson, H. (1999). Asymmetric time series and temporal aggregation. *The Review of Economics and Statistics*, 81, 341-344.

Cachon, G., & Fisher, M. (1997). Campbell Soup's continuous replenishment program: evaluation and enhanced inventory decision rules. *Production and Operations Management*, 6, 266-276.

Cecere, L. (2012). Big data: go big or go home? *Supply Chain Insights*, July.

Chen, L., & Lee, H.L. (2012). Bullwhip effect measurement and its implications. *Operations Research,* 60, 771-784.

Cognizant, (2012). Big data's impact on the data supply chain. *20-20 Insights*, May.

Das, C., & Tyagi, R. (1999). Effect of correlated demands on safety stock centralization: patterns of correlation versus degree of centralization. *Journal of Business Logistics,* 20, 205-214.

DeHoratius, N., & Raman, A. (2008). Inventory record inaccuracy: An empirical analysis. *Management Science*, 54, 627-641.

East, R., Lomax, W., Willson, G., & Harris, P. (1994). Decision making and habit in shopping times. *European Journal of Marketing,* 28, 56-71.

Evers, P.T. (1995). Expanding the square root law: an analysis of both safety and cycle stocks. *Logistics & Transportation Review,* 31, 1.

Evers, P.T. (1996). The impact of transshipments on safety stock requirements. *Journal of Business Logistics,* 17, 109-133.

Evers, P.T. (1997). Hidden benefits of emergency transshipments. *Journal of Business Logistics,* 18, 55-76.

Evers, P.T., & Beier, F.J. (1993). The portfolio effect and multiple consolidation points: a critical assessment of the square root law. *Journal of Business Logistics,* 14, 109-126.

Evers, P.T., & Beier, F.J. (1998). Operational aspects of inventory consolidation decision making. *Journal of Business Logistics,* 19, 173-189.

Finn, R. (2004). Weekly vs monthly forecasting in the supply chain. *Logistics and Transport Focus,* 6, 22-28.

Gilbert, K. (2005). An ARIMA supply chain model. *Management Science,* 51, 305-310.

Goldsby, T.J., Griffis, S.E., & Roath, A.S. (2006). Modeling lean, agile and leagile supply chain strategies. *Journal of Business Logistics*, 27, 57-80.

Granger, C.W.J., & Siklos, P.R. (1995). Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration: theory and evidence. *Journal of Econometrics,* 66, 357-369.

Hendry, D.F. (1992). An econometric analysis of TV advertising expenditure in the United Kingdom. *Journal of Policy Modeling*, 14, 281-311.

Hofer, C., Jin, H., Swanson, D., Waller, M.A., & Williams, B. (2012). The impact of key retail accounts on supplier performance: a collaborative perspective of resource dependency theory. *Journal of Retailing*, 88, 412-420.

Lapide, L. (2007). Sales and operations planning mindsets. *Journal of Business Forecasting*, 26, 21-31.

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

Liao, H., & Chuang, A. (2004). A multilevel investigation of factors influencing employee service performance and customer outcomes. *Academy of Management Journal,* 47, 466-485.

Lutkepohl, H. (1987). *Forecasting Aggregated Vector ARMA Processes,* Berlin: Springer-Verlag.

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: results of a forecasting competition. *Journal of Forecasting,* 1, 111-153.

Makridakis, S., & Hibon, M. (2000). The M3-competition: results, conclusions and implications. International *Journal of Forecasting,* 16, 451-476.

Marcellino, M. (1999). Some consequences of temporal aggregation in empirical analysis. *Journal of Business and Economic Statistics,* 17, 129-136.

Mahmoud, M.M. (1992). Optimal inventory consolidation schemes: a portfolio effect analysis. *Journal of Business Logistics*, 13, 193-214.

McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard Business Review*, October.

McCarthy, T.M., Davis, D.F., Golicic, S.L., & Mentzer, J.T. (2006). The evolution of sales forecasting management: a 20-year longitudinal study of forecasting practices. *Journal of Forecasting*, 25, 303-324.

Mentzer, J.T., & Cox, J.E. (1984). A model of the determinants of achieved forecast accuracy. *Journal of Business Logistics,* 5, 143-155.

Mentzer, J.T., & Kahn, K.B., (1995). Forecasting technique familiarity, satisfaction, usage, and application. *Journal of Forecasting,* 14, 465-476.

Nijs, V.R., Srinivasan, S., & Pauwels, K. (2007). Retail-price drivers and retailer profits. *Marketing Science*, 26, 473-487.

Nijman, T.E., & Palm, F.C. (1990). Predictive accuracy gain from disaggregate sampling in ARIMA models. *Journal of Business and Economic Statistics,* 8, 405–415.

Nikolopoulos, K., Syntetos, A.A., Boylan, J.E., Petropoulous, F., & Assimakopoulos, V. (2011). An aggregate-disaggregate intermittent demand approach (ADIDA) to forecasting: an empirical proposition and analysis. *Journal of the Operational Research Society,* 62, 544-554.

Olivia, R., & Watson, N. (2011). Cross-functional alignment in supply chain planning: a case study of sales and operations planning. *Journal of Operations Management,* 29, 434-448.

Orszag, P. (2012). End-of-month hunger hurts students on food stamps. Retrieved from http://www.bloomberg.com/news/2012-04-03/end-of-month-hunger-hurts-students-on-food-stamps.html.

Pauwels, K., Currim, I., Dekimpe, M.G., Ghysels, E., Hanssens, D.M., Mizik, N., & Naik, P. (2004). Modeling marketing dynamics by time series econometrics. *Marketing Letters,* 15, 167-183.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical Linear Models.* Sage Publishing, Thousand Oaks, CA.

Ravichandran, T., & Liu, Y. (2011). Environmental factors, managerial processes, and information technology investment strategies. *Decisions Sciences,* 42, 537-574.

Rexhausen, D., Pibernik, R., & Kaiser, G. (2012). Customer-facing supply chain practices – the impact of demand and distribution management on supply chain success. *Journal of Operations Management*, 30, 269-281.

Ronen, D. (1990). Inventory centralization: the square root law revisited again. *Journal of Business Logistics*, 11, 129-138.

Rossana, R.J., & Seater, J.J. (1995). Temporal aggregation and economic time series. *Journal of Business & Economic Statistics,* 13, 441-451.

Sbrana, G., & Silvestrini, A. (2013). Aggregation of exponential smoothing processes with an application to portfolio risk evaluation. *Journal of Banking and Finance,* 37, 1437-50.

Schmarzo, B. (2012). Big data in traditional retail – Part I", *InFocus*, Retrieved from https://infocus.emc.com/william_schmarzo/big-data-in-traditional-retail-part-i/.

Schoenherr, T. & Swink, M. (2012). Revisiting the arcs of integration: cross-validations and extensions. *Journal of Operations Management,* 30, 99-115.

Silvestrini, A., & Veredas, D. (2008). Temporal aggregation of univariate and multivariate time series models: a survey. *Journal of Economic Surveys,* 22, 458-497.

Sims, C.A. (1971). Discrete approximations to continuous time distributed lags in econometrics. *Econometrica,* 29, 545-563.

Tallon, W.J. (1993). The impact of inventory centralization on aggregate safety stock: the variable supply lead time case. *Journal of Business Logistics,* 14, 185-204.

Tiao, G.C. (1972). Asymptotic 88ehavior of temporal aggregates of time series. *Biometrika,* 59, 525-531.

Vogt, J.J. (2010). The successful cross-dock based supply chain. *Journal of Business Logistics,* 31, 99-119.

Wei, W.W.S. (1978). Some consequences of temporal aggregation in seasonal time series models. *Seasonal Analysis of Economic Time Series,* Washington, D.C., U.S. Department of Commerce, Bureau of the Census.

Wei, W.W.S. (1982). The effects of systematic sampling and temporal aggregation on causality – a cautionary note. *Journal of the American Statistical Association,* 77, 316-319.

Williams, B.D., & Waller, M.A. (2010). Creating order forecasts: point-of-sale or order history? *Journal of Business Logistics,* 31, 231-251.

Williams, B.D., & Waller, M.A. (2011). Top-down vs. bottom-up demand forecasts: the value of shared point-of-sale data in the retail supply chain. *Journal of Business Logistics,* 32, 17-26.

Zinn, W., & Bowersox, D.J. (1988). Planning physical distribution with the principle of postponement. *Journal of Business Logistics,* 9, 117-136.

Zinn, W., Levy, M., & Bowersox, D.J. (1989). Measuring the effect of inventory centralization/decentralization on aggregate safety stock: the ‘square root law’ revisited. *Journal of Business Logistics*, 10, 1-14.

Zotteri, G., & Kalchschmidt, M. (2007). Forecasting practices: empirical evidence and a framework for research. *International Journal of Production Economics*, 108, 84-99.

Table 1 – MAD by category, temporal aggregation, and demand signal

	Temporal Aggregation	
	Weekly	Tagg4
Category 1		
Order	66.161	61.240
POS	56.334	59.384
Category 2		
Order	331.690	322.460
POS	325.954	349.450

Table 2 – Hierarchical linear modeling results

Variable		Null	Model 1	Model 2	Model 3
Fixed effects					
	Intercept	142.97 (35.41)***	23.88 (17.97)	18.78 (18.00)	32.53 (18.08)*
β_1	*Mean*		0.35 (0.03)***	0.35 (0.03)***	0.35 (0.03)***
β_2	*AR*		-3.53 (2.62)	-2.05 (2.61)	-11.09 (4.90)**
β_3	*Agg*		0.79 (0.68)	9.14 (0.96)***	-30.47 (1.38)***
β_4	*POS*		-5.67 (2.44)**	-15.32 (2.56)***	13.23 (4.29)***
β_5	*Agg x POS*			16.66 (1.36)***	-26.35 (5.00)***
β_6	*POS x AR*				-27.86 (7.57)***
β_7	*Agg x AR*				54.80 (5.64)***
β_8	*Agg x POS x AR*				35.56 (8.62)***
Random effects					
	CAT	16648.14 (6624.82)**	2080.27 (1343.01)	2089.04 (1347.71)	2148.16 (1374.25)
	DC	4913.48 (844.04)***	3362.99 (615.60)***	3362.95 (615.64)***	3357.18 (614.36)***
	FP	326.18 (56.80)***	326.26 (56.80)***	326.31 (56.80)***	326.35 (56.80)***
	Residual	2995.91 (26.57)***	2975.92 (26.39)***	2958.50 (26.24)***	2948.89 (26.15)***
-2 Log Likelihood		278204.13	277982.05	277832.8	277750.171
			222.08	149.25	82.629

Notes: Standard errors are shown in parentheses.
*p<0.1, **p<0.05, ***p<0.01.

Figure 1 – Quasi-experimental design

	Weekly	Monthly
Order	AGG = 0 POS = 0 N = 6,396	AGG = 1 POS = 0 N = 6,396
POS	AGG = 0 POS = 1 N = 6,396	AGG = 1 POS = 1 N = 6,396

Figure 2 – MAD and autocorrelation factor

	Weekly	Monthly
Order $AR_{Order} = -0.18$	150.75 (148.97)	143.23 (159.42)
POS $AR_{POS} = 0.73$	133.56 (148.97)	142.75 (162.43)

Figure 3: Interaction plot of temporal aggregation and retail demand signal

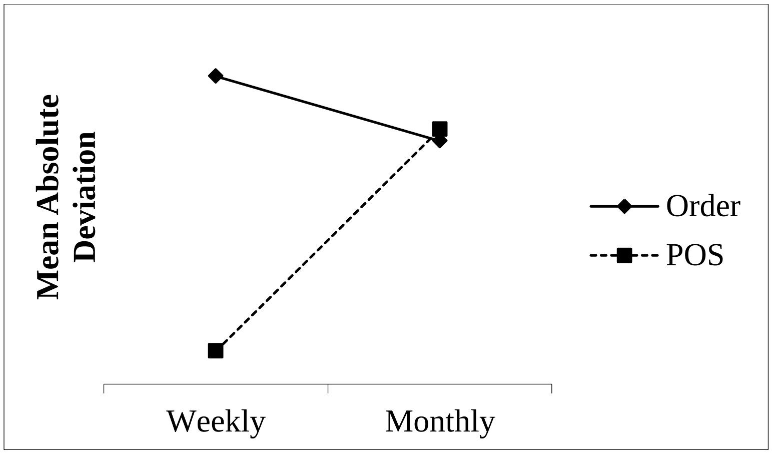


Figure 4 – Autocorrelation's impact on MAD for weekly and monthly POS demand

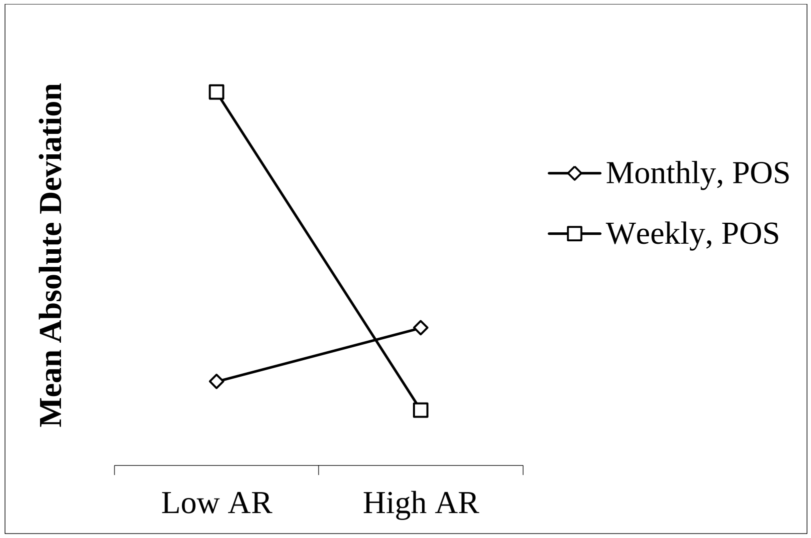


Figure 5 – Autocorrelation's impact on MAD for weekly and monthly DC demand

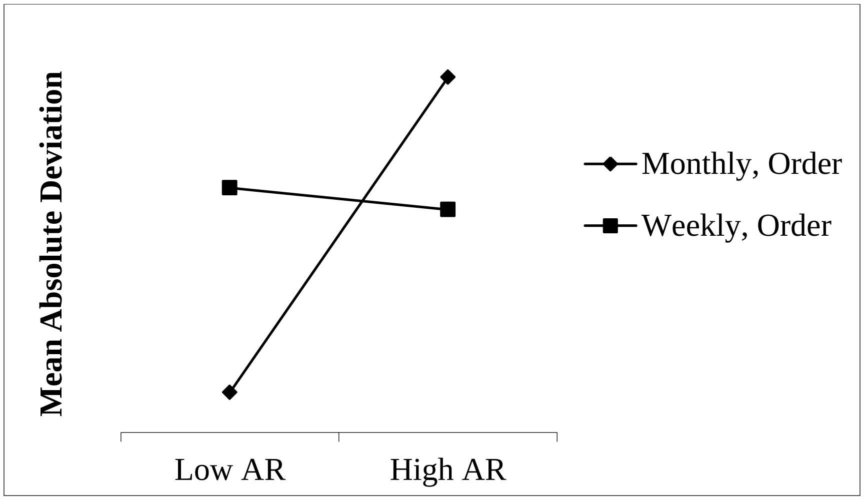
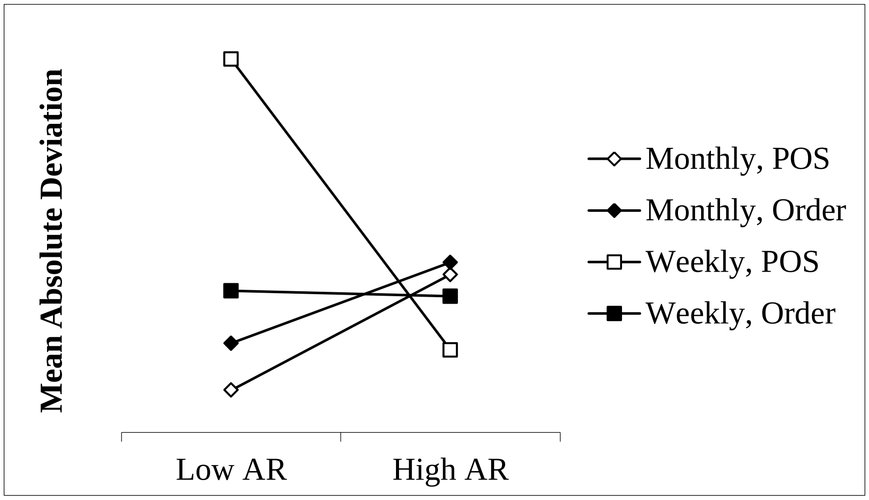


Figure 6 – Autocorrelation's impact on MAD for weekly and monthly POS and DC demand



Chapter 5 – DEMAND SIGNAL AND MODEL SELECTION FOR SEASONAL FORECASTING—THE MODERATING ROLE OF BULLWHIP

V. DEMAND SIGNAL AND MODEL SELECTION FOR SEASONAL FORECASTING—THE MODERATING ROLE OF BULLWHIP

A. INTRODUCTION

Bullwhip is defined as the amplification of demand variance as demand signal travels upstream along the supply chain (Lee et al., 1997) to obscure visibility to true demand. Increased demand variance can command higher than necessary inventory due to poor customer demand forecast (Agarwal & Holt, 2005) and lead to significant write-downs (Hanssens, 1998). Although conventional literature prescribes information sharing by downstream firms as a key remedy to assist suppliers with demand planning (Lapide, 1999; Lee et al., 1997; Cachon & Fisher, 2000), the actual value of this practice for forecasting purposes had come under doubt (Williams & Waller, 2010).

Demand signals shared by firms downstream along the supply chain are relatively free of idiosyncratic distortions. With a more accurate view of consumer demand, suppliers are believed to be able to reduce uncertainty in the supply chain (Cachon & Fisher, 2000; Lee et al., 2000). However, in forecasting short term customer demand, suppliers are not necessarily preoccupied with predicting consumer demand. Instead, their immediate customers are the retailers, whose ordering patterns may or may not be directly in response to short term consumer demand (Parkany, 1961). Determining the appropriate forecast input is particularly important for seasonal products for both suppliers and retailers. Compressed selling season means that too much inventory leads to increased discounts while too little results in lost sales.

Because of their significant spike in short-term demand, seasonal products command greater flexible transportation, storage, and labor capacities. In response, retailers often engage in ordering patterns that deviate from consumer demand in an attempt to smooth seasonal demand spikes, thereby alleviating retail operations constraints (Bray & Mendelson, 2012). Thus, shared consumer demand signal is not necessarily the most appropriate forecast input for retail orders. Along the lines of this principle, Williams and Waller (2010) found that POS data may not always be the optimal forecast input for suppliers. However, bullwhip increases both variance components of a demand series. Therefore its influence on the use of downstream demand signal for customer demand forecasting warrants investigation.

In addition to demand signal selection, model selection is equally ambiguous in seasonal forecasting. For time series data which are known a priori to be seasonal, the seasonal effect may be either additive or multiplicative (Chatfield, 1978). Although theoretical literature indicates an overall preference for the multiplicative model (Chatfield & Yar, 1988), there is little empirical evidence to provide guidance for the model selection process. Furthermore, most software packages utilized by companies today do not offer diagnostics to assist planners with model selection. Considering that bullwhip alters the cyclicality and variance properties of demand signals (e.g., Thornhill & Naim, 2006), its effect on the seasonal forecast model selection warrants exploration as well.

Using a large sample of demand signals, point-of-sale (POS) and order data, for products from a high volume seasonal category, this study contributes to the body of literature on data science in supply chain management in two folds. First, our findings corroborate with Williams and Waller (2010), POS is not appropriate for suppliers to use for demand planning. However, bullwhip's distortionary effect closes the forecast performance gap between POS and order. Second, our exploratory analysis in forecast model preference reveals that when POS is the demand signal utilized, multiplicative generally outperforms the additive forecast model. But as bullwhip increases, the additive model begins to outperform multiplicative model.

B. Literature and Hypotheses

<u>Demand Signal and Bullwhip</u>

Recent advancement in supply chain management strategies gave rise to collaborative forecasting practices that place a particular emphasis on information sharing as firms positions downstream along the supply chain share their observed demand with their partners upstream (e.g., Cachon & Fisher, 2000). By sharing relevant and meaningful information (Kaipia & Hartiala, 2006), companies may improve both inter-functional and inter-company linkages (Schoenherr & Swink, 2012), empower collaboration capabilities (Allred et al., 2011), and tame the bullwhip effect (Lee et al., 1997) to generate superior statistical forecasts (Rexhausen et al., 2012). In turn, superior statistical forecasts may be used to synchronize supply chain activities across both the company (Olivia & Watson, 2011) the supply chain (Cao & Zhang, 2011).

In the retail supply chain, suppliers may either replenish stores directly or through a retailer's distribution centers (DC). Often, large retailers follow a centralized distribution process where retail stores are replenished by regional DCs. Electronically-transmitted point-of-sale (POS) data along with retail orders are then aggregated at the DC level. Based on data collected from retail stores, DCs place periodic orders with the suppliers. These orders, in turn, become suppliers' most readily observed customer demand and often serve as the principle demand signal suppliers utilize for forecasting future customer demand.

Based on the myriad advantages associated with information sharing, a logical conclusion may be drawn that downstream demand signals are always superior. Yet, recent research indicates that POS data does not always outperform DC orders in forecast accuracy (Williams & Waller, 2010). As retail sales accumulate at the store level, replenishment and execution processes at various nodes of a retailer's internal distribution network (Vogt, 2010) influence DC orders placed with its suppliers. In addition, idiosyncratic effects such as warehouse management systems (Autry et al., 2005), and supply chain processes such as postponement (Zinn & Bowersox, 1988) and inventory management policies (Evers & Beier, 1998; Goldsby et al., 2006) may all impact retail orders.

Retailers frequently utilize inventory and order management strategies to mitigate short-term demand variability and alleviate operations planning difficulties (Dooley et al., 2010). For seasonal products in particular, retailers can reduce operational strains placed on its distribution network through smoothing spikes in seasonal demand (Cachon et al., 2007; Bray & Mendelson, 2012). A frequent result of such demand signal processing is altered cyclicality, because the retailer tends to steadily build inventory during low demand seasons for rapid depletion when seasonal demand peaks. Thus, both historic and future DC orders associated with seasonal products tend to reflect the retailer's systematic ordering patterns and less so consumer demand.

Under collaborative forecasting practices, suppliers often have access to both demand signals concurrently. Whereas the POS offers a more accurate view of consumer demand, order data possesses demand variance entails retail inventory and order management policies. Thus, while the supplier may leverage POS data to forecast future consumer demand and gain market insight in the long term, its immediate short term concerns remain fulfilling retail customer demand in the near term. Since future DC orders will likely continue to follow past ordering patterns, we expect the use of POS data as the forecast input to increase forecast error.

> H1: The use of POS demand signal as the forecast input is positively associated with forecast error.

The bullwhip effect (e.g., Lee et al., 1997) is a major contributor to forecast error. The reason is simple and compelling: additional demand variance introduced by managerial and behavioral idiosyncrasies result in unpredictable alteration to demand pattern. However, Bray and Mendelson (2012) identified two main components of bullwhip in seasonal and stochastic variance. Whereas the seasonal smoothing effect lowers the magnitude of the bullwhip effect (Cachon et al., 2007; Bray & Mendelson, 2012), stochastic influences amplify it.

In addition to retail ordering and inventory management policies, DC orders also reflect idiosyncratic effects such as behavioral factors (Lee et al., 1997; Kaipia et al., 2006), which tend to amplify demand variance. While changes due to seasonal inventory management policies can be expected to recur with each cycle, behavioral factors are far less predictable. Moreover, seasonal and stochastic variance components are hard to distinguish and even more difficult to separate. Further, whereas the seasonal component of demand variance can be interpreted as constant or a function of mean demand (Chatfield & Yar, 1988), unpredictable behaviors' inflates demand variance at random.

As characterized by Bray and Mendelson (2012), the "tug-of-war" between seasonal smoothing and stochastic influence often results in net increase in demand variability upstream along the supply chain. Heightened bullwhip reflects greater influence from stochastic amplification on order variability over deliberate and recurring seasonal smoothing. As a result of increased bullwhip, forecast accuracy deteriorates. Forecasters may mitigate the detrimental effect of high bullwhip on forecast accuracy by utilizing downstream demand signals (Lee et al., 2000). Thus, the use of POS data for seasonal products that exhibit high degree of bullwhip should lead to more accurate customer demand forecast. Hence:

> H2: The use of POS demand signal as the forecast input negatively moderates the positive effect of bullwhip on customer demand forecast error.

C. Methodology

<u>Demand Forecast Competition</u>

To test our hypotheses, we develop a demand forecast competition (DF-Competition) by using monthly data collected by a large national retailer. The design of this competition builds on Williams and Waller (2011), which develops customer demand forecast utilizing two sources of demand signals, namely POS and DC orders. POS is defined as the cumulative number of units sold during each month at all stores served by a specific DC; each DC may serve up to approximately one-hundred stores. All DCs experience similar volumes of aggregate retail demand. Additionally, an order is defined as the cumulative number of units ordered by the particular retailer DC during the same month. Since the particular retailer in our sample only purchases each SKU by cases, POS and orders are therefore measured in cases as well.

To generate customer demand forecast, we utilize two types of commonly-used seasonal models: Holt-Winter's additive and multiplicative. The additive model assumes that the seasonal effect is constant over time. The multiplicative model assumes that the seasonal effects are proportional to the deseasonalized mean level. Note that the deseasonalized mean level may be modified by an additive trend term (Chatfield, 1978). While additive and multiplicative models provide distinct treatment to calculate smoothed components, the additive model procedure for estimating future demand can be described by:

$$\hat{Y}_{t+m} = (\mu_t + \tau_t m) + S_{t+m}$$

where $\hat{Y}$ and *S* are the estimated demand and smoothed seasonal factor at time *t*, for *m* periods into the future, respectively. μ and τ are the smoothed level demand and trend at time t. According to McKenzie (1976), the additive model is optimal for only a particular ARIMA process, therefore it is not considered to be a flexible description of possible seasonal processes. The multiplicative model procedure can be described by:

$$\hat{Y}_{t+m} = (\mu_t + \tau_t m) S_{t+m}$$

The smoothed components for each forecast model must be estimated with three smoothing parameters (alpha, beta, & gamma). To ensure the level of rigor in our forecast competition, we utilize twenty-seven different combinations of the three smoothing parameters. The specific levels were chosen based on extant forecast literature (Silver, Pyke, & Peterson 1998), which prescribes the recommended maximum, optimal, and minimum levels for each parameter. The specific levels can be seen in Appendix I. For each SKU-DC combination, we generate separate forecasts based on each combination of smoothing parameters for both additive and multiplicative models.

First, we utilize two years of monthly observation for in-sample model estimation to obtain components of level, trend, and seasonal components of demand forecast. Next, we compare the monthly forecast with out-of-sample monthly demand to measure forecast performance. Forecast performance is measured as mean absolute percentage error (MAPE), which is a commonly-used forecast error metric (McCarthy et al. 2006; Mentzer & Kahn 1995). To construct the additive versus multiplicative forecast experiment, we create a two-by-two forecast design to reflect common supplier forecast settings, resulting in four distinct groups.

Figure 1 illustrates our DF-competition setup. We estimate for each SKU-DC combination the smoothed components for level, trend, and seasonality for order and POS data using both additive and multiplicative Holt-Winters seasonal forecasting models. We next utilize the estimated forecast components to generate DC orders for each SKU-DC combination over a 6 months out-of-sample forecast horizon. Next, we calculate for each SKU-DC combination their DC order forecast error using estimated demand from POS and order data.

Model Specification

Demand forecast requires forecasters to first consider demand variance (D) and forecasting method (M). Therefore, forecast error (MAPE) is a basic function of D and M:

MAPE = f(D,M)

A common measure of demand uncertainty is simply the variance of demand. However, variance alone cannot adequately inform probability distributions. Dekimpe and Hanssens (1995) conducted a meta-analysis of 44 studies that include 180 sales series and found that 68% of the sales series are nonstationary. Thus, a trend component should be included to account for change in mean demand. Ideally, demand variance for seasonal products should be separated as different components reflecting exogenous factors that induce demand variability such as calendar, length of seasonal cycle, and in-season factors such as weather and holidays. However, many of these variance components are stochastic and cannot be estimated. Therefore, we define demand variance (D) as a function of the variance (Var) and trend (Trend):

D = f(Var,Trend)

Our demand forecast is generated with two commonly-utilized seasonal forecasting models—Holt-Winter's *additive* and *multiplicative* models. Both models incorporate a set of three smoothing parameters to place desired emphasis on level, trend, and seasonality for generating new statistical forecasts. Therefore the smoothing parameters *alpha*, *beta*, and *gamma*, all leverage unique influence on the demand forecast. Once the model and parameters are selected, forecasters need to further determine the forecast horizon (*Horizon*). Thus, we define forecasting method as the function below:

M = f(Additive, Multiplicative, alpha, beta, gamma, Horizon)

Therefore, we define the forecast error of seasonal products as:

MAPE = f(Var, Trend, Additive, Multiplicative, alpha, beta, gamma, Horizon).

Although the above function is complete for a typical seasonal forecasting setting, this study tests the effect of the demand signal utilized in seasonal forecasting as well. Specifically, the forecast experiment pits two sources of demand signals in a DF-Competition from the perspective of a supplier. One demand signal is the retail order history (*Order*). The second demand signal is the POS history (*POS*) that is visible to suppliers in a collaborative forecast setting. A primary statistical distinction between the two demand signals is that retail order history is more susceptible to the bullwhip effect. Amplified variance (i.e., bullwhip) induces randomness to adversely impact forecast performance. Therefore, we further add *Bullwhip* to the function as outlined above:

MAPE = f(var, trend, Additive, Multiplicative, alpha, beta, gamma, Horizon, POS, Order, Bullwhip)

In transforming the above definition of forecast error in the context of this study to a testable model, we make the following adjustments. First, initial tests (Durbin-Watson = 2.81) indicated that the dependent variable possesses significant first order autocorrelation, therefore we include its lagged term to account for potential biases. Second, most seasonal factors are assumed to be multiplicative (Chatfield & Yar, 1988). Therefore, we code *Additive* as a binary variable to account for its influence, for which 1 indicates a forecast error that is generated with the additive model and 0 if multiplicative. Third, a focal variable of this study is the effect of utilizing shared consumer demand signal—POS, from the retailer (H1). We code *POS* as a binary variable as well. Hence, we drop *Multiplicative* and *Order* in our regression equation below. Lastly, H2 concerns the moderating influence of *Bullwhip* on the effect of *POS*. Thus, an interaction term, *Bullwhip*POS*, is included in our full model, presented below:

$$MAPE_{idm} = \beta_1 Var_{idm} + \beta_2 Trend_{idm} + \beta_3 Horizon_{idm} + \beta_4 alpha_{idm} + \beta_5 beta_{idm} + \beta_6 gamma_{idm} + \beta_8 MAPE_{idm-1} + \beta_7 Additive_{idm} + \beta_9 Bullwhip_{id} + \beta_{10} POS_{idm} + \beta_{11} Bullwhip_{id} * POS_{idm} + \varepsilon_{idm}$$

Where subscripts *i*, *d*, and *m* designate the product (*i* = 1, …, 6), DC (*d* = 1, …, 6), and time period (*m* = 1, …, 6), respectively.

Data Collection and Measures

Our data comes from a high volume and highly seasonal grocery category that is one of the most commonly-shopped categories in a typical grocery retail format. Specifically, our sample includes two-and-a-half years of monthly data for six SKUs from six regional U.S. retail DCs. These DCs are operated by one of the supplier's largest retail customers. In all, our sample contains thirty-six unique SKU-DC combinations of thirty monthly DC order and POS sales series. DC orders are defined as the total monthly cases of a product ordered by a particular DC to the supplier while POS is the cumulative monthly sales of the retail stores served by the particular DC.

The objective of our DF-Competition is to compare the customer demand forecast errors, from the supplier's perspective, based on combinations of demand signal and forecast model utilized. Of the thirty monthly demand observations, we utilize the first twenty-four months for in-sample estimation of forecast parameters, which are then used to forecast demand and calculate out-of-sample forecast errors for the remaining six months.

Dependent variable and variables of interest

As mentioned previously, forecast error is measured as MAPE. It is also our dependent variable for the regression model. MAPE measures forecast error by averaging the absolute value of the percent error for each forecast. It is calculated by first taking the absolute value of the actual monthly orders (A) less the order forecast ($\hat{F}$). Next, the absolute value of the difference is divided by A. Finally, the average of the cumulative total percent error is calculated at each forecast horizon. The calculation for $MAPE_{idm}$ for customer demand forecast is shown:

$$MAPE_{idm} = \left(\frac{1}{n}\right) \sum_{m=1}^{n} \left| \frac{(A_{idm} - \hat{F}_{idm})}{A_{idm}} \right|$$

In addition, our variables of interest include $Additive_{idm}$, POS_{idm}, and $Bullwhip_{idm}$. $Additive_{idm}$ and POS_{idm} are both binary variables so no further calculation is needed. The bullwhip effect is consistently defined as the amplification of demand variability due to managerial and behavioral activities such as demand signal processing (e.g., Lee et al., 1997; Sterman, 1989; Dejonckheere et al., 2003; Chen & Lee, 2012). Thus, following the established definition, $Bullwhip_{idm}$ is calculated as:

$$Bullwhip_{id} = \frac{Var(O_{id})}{Var(D_{id})}$$

Control variables

Our control variables include Var_{idm} and $Trend_{idm}$, which are specific to each SKU-DC combination, and forecasting parameters $alpha_{idm}$, $beta_{idm}$, $gamma_{idm}$, and $Horizon_{idm}$. Because the retailer processes demand signals from the consumers to generate orders, therefore Var_{idm} is calculated simply as the variance of POS for each SKU-DC combination over the in-sample estimation period of twenty-four months. Similarly, $Trend_{idm}$ is also calculated for the in-sample estimation period by simply regressing monthly POS for each SKU-DC combination against time, i.e., $Order_{idm} = \gamma_1 t_{idm} + \varepsilon_{idm}$, in which $\gamma_1$is the linear trend coefficient for each time series. Recall that we utilize three different values for each of the forecast parameters $alpha_{idm}$, $beta_{idm}$, and $gamma_{idm}$ (Appendix I). They are included as continuous variables to control for potential systematic influence on forecast error. Lastly, $Horizon_{idm}$ is the forecast horizon associated with each forecast, and it ranges from 1 to 6.

D. Descriptive Statistics

We present, in Table 1, the descriptive statistics and correlations for MAPE along with the various measures of demand distribution characteristics. The overall average MAPE over a six period forecast horizon is approximately 78%, which once again highlights the difficulty associated with forecasting seasonal orders. The median MAPE is about 49%, indicating that the distribution of MAPE is somewhat skewed by a small number of observations that have large magnitude of forecast error. Our sample demonstrates substantial bullwhip effect with the average variance ratio being 1.27. In addition, demand for SKUs in this study demonstrated a positive trend, with 19.27, during our data collection period

Figure 2 shows comparison of MAPE by forecasting method, for each demand signal. As expected, the multiplicative forecast model significantly outperforms the additive model for both demand signals by approximately thirty percentage points. While Figure 2 suggests that there is no difference between the two demand signals with respect to average MAPE, Table 2 presents a more detailed view. We first segment MAPE by forecast models and then by demand signal to compare the forecast performance of demand signals within each model. Despite the relatively similar average MAPE, POS outperforms order approximately 62% of the time. However, when order outperforms, the average improvement to forecast error is much higher. Therefore, demand signal input can significantly impact forecast performance for both models.

In addition, we also segment MAPE by demand signal and then by forecast model to compare difference between models within each demand signal. Pair-wise comparison revealed that while the multiplicative model outperforms additive 56% of the time, the resultant improvement to MAPE is 75 percentage points when the demand signal utilized is POS, and 62 percentage points for order. For the other 44% of pair-wise comparisons when additive outperforms multiplicative, their difference in MAPE is 18 percentages points for when forecast input is POS, and 13 percentages points for order. Thus, while multiplicative outperforms additive a small majority of the time, the resulting improvement is much higher than when the reverse is true. On the other hand, additive still outperforms multiplicative on many occasions, and yield sizeable improvement to forecast error. To better assist forecasters with model selection, we conduct an exploratory analysis to examine the influence of demand characteristics on model superiority after we test our hypotheses.

E. Hypothesis Testing and Results

We hierarchically enter our variables into our model. First, all control variables are entered into Model 1 (Table 3). Coefficients for all control variables are significant at the 0.01 level and are of the expected signs. Durbin-Watson statistic is 2.16, which alleviates autocorrelation concerns with MAPE. The control variables collectively explain 32.9% of the total variance in MAPE.

Next, we enter our direct effects of interest. Total variance explained increased to 33.6%, while Akaike's Information Criterion decreased from 3.752 to 3.742. Results show that bullwhip is positively associated with MAPE ($\beta_9 = 0.429, p < 0.01$), which supports the notion that demand signal distortion leads to less accurate forecast. H1 argues that for seasonal products, the use of POS increases customer demand forecast error. Statistical evidence supports H1 ($\beta_{10} = 0.211, p < 0.01$), indicating that the use of POS is positively associated with MAPE. In Model 3, we enter the interaction term for POS and Bullwhip. Results show a significant interaction ($\beta_{10} = -0.963, p < 0.01$) to lend support for H2. The use of POS data as forecast input can mitigate bullwhip's inflationary effect on forecast error.

Lastly, models indicate that additive positively influences MAPE ($\beta_8 = 0.345, p < 0.01$), which supports the belief that multiplicative, in general, is the preferred seasonal forecast model (Chatfield & Yar, 1988). However, recall in Table 2 that the additive model does outperform the multiplicative 44% of the time, often to sizeable improvements between 13 to 18 percentage points. Clearly, a generalized statement in broad support of multiplicative over additive is not appropriate. Together, the descriptive statistics and regression results further call an exploratory analysis to examine determinants of model choice between the two Holt-Winters seasonal forecast models.

Model Robustness

Although we controlled for demand characteristics associated with each SKU, as well as forecasting parameters, there may be additional idiosyncratic difference exist among different SKUs due to consumer preference. Thus, to verify that these potential idiosyncratic effects do not adversely impact conclusions that may be drawn from our statistical model, we performed an additional model (Alt. Model, Table 3) to include product fixed effects. As shown, all parameters estimates of interest remained qualitatively the same, with the lone exception of the direct effect of bullwhip is no longer significant with the inclusion of its interaction effect with POS. However, this is expected since the bullwhip ratio in this study is operationalized at the product level.

F. Exploratory Analysis

Seasonal Forecasting Model Selection

While most seasonal factors are considered multiplicative in nature (Chatfield & Yar, 1988), our initial comparison of between-model forecast error indicates that additive outperforms multiplicative 44% of the time. Considering that the multiplicative is most commonly applied, substantial opportunities remain in increasing seasonal forecast accuracy by diagnosing demand characteristics that can be used to make the optimal model choice.

<u>Exploratory Logistic Regression</u>

First, we code AMM_{idm} as a binary variable that represents when the multiplicative model outperforms the additive ($AMM_{idm} = 1$) for SKU *I*, DC *d*, and time *m*. We utilize all demand characteristics from the regression model, Var_{idm}, $Trend_{idm}$, $Horizon_{idm}$, and $Bullwhip_{idm}$. In addition, we include the binary variable, POS_{idm}, to indicate when POS is utilized as the demand signal for forecasting customer demand. Finally, to perform a thorough examination of how the demand characteristics of different demand signals influence model selection, we also include all of POS_{idm}'s two-way interaction with demand characteristics.

Since our dependent variable, AMM_{idm}, is a binary variable, parameters should not be estimated with OLS (Greene, 2011). Therefore we estimate our model with logistic regression to obtain the change in the probability that our dependent variable is 1, with change in each independent variable. We enter the direct and interaction effects in blocks. In Table 4, we present the logistic regression results.

We enter the direct and interaction effects hierarchically. Direct effects (Model 1, Table 4) suggest that variance, forecast horizon, bullwhip, and the use of POS demand signal all result in increased likelihood of the multiplicative model outperforming the additive model. In addition, trend is the only direct effect that decreases the likelihood of the multiplicative outperforming the additive, which suggests that the additive model's relatively conservative treatment of the seasonal factor may be preferred for demand forecasts farther into the future.

The interaction effects of logistic regression (Model 2, Table 4) provides a more detailed examination of seasonal forecasting model selection given two demand signals. All interaction terms are between POS and the four demand distribution characteristics variables. The McFadden R-Squared increased to 47.5% to match a drop in both the negative log of likelihood and Akaike's Information Criterion. We find that POS negatively moderates bullwhip and trend in their respective influence on the likelihood of multiplicative outperforming additive. In addition, POS positively moderates variance and horizon in their respective influence.

G. Discussion and Implications

In the retail supply chain, retailers engage suppliers in collaborative demand planning by sharing POS data, which are demand signals observed at the consumer level. While information sharing is only one part of a broader supply chain integration effort, it has significant impact on operational, and thereby, financial performance (Germain & Iyer, 2006). The value in utilizing POS data for forecasting customer demand is well recognized (e.g., Williams & Waller, 2010). However, POS's value in seasonal forecasting is less clear. In particular, POS and order data have unique information. Whereas the former more closely reflects consumer demand, order data contains potential indicators of retailer inventory management policies. Further complicating seasonal forecasting is the existence of two competing models, namely Holt-Winters multiplicative and additive model. We attempt to diagnose the effect of demand distortion and selection on both forecast performance and model choice.

Results from this study corroborate with Williams and Waller (2010), and show that using POS for forecasting seasonal customer demand can increase forecast error. That is not to say that there is no inherent value in the information provided by POS. Rather, a primary benefit of POS data is visibility to consumer demand without demand distortion. Due to retailers' propensity to smooth seasonal consumer demand (Bray & Mendelson, 2012), future retail orders for seasonal products are more likely to follow past order patterns. POS therefore is not the optimal demand signal for seasonal forecasting, because it reflects only consumer demand and not retail ordering policies.

However, the performance gap between POS and order decreases as the degree of bullwhip increases. As illustrated in Figure 2, order-based forecast accuracy decreases as bullwhip increases, while the opposite is true for when forecast utilizes POS data. As retailers engage in seasonal smoothing policies, “rhythmic” ordering can result in cyclical properties (e.g., McCullen & Towill, 2002) that future orders likely follow (Parkany, 1961). From a supplier’s perspective, POS provides clarity to consumer demand patterns but not insight into retail ordering policies. With increased bullwhip, stochastic variance component of order data becomes amplified due to unpredictable behavioral factors to negatively impact forecast accuracy in two ways. First, unpredictable behavioral factors generally do not follow any statistical patterns. Second, inflated variance has a destabilizing effect on the overall demand forecast. Thus, when bullwhip is low, order history may allow suppliers to anticipate future customer orders by incorporating statistical properties due to retail ordering policies. However, as bullwhip increases, the resultant noise may obscure actionable intelligence derived from order history. Thus, while order may be preferable to POS in seasonal forecast, bullwhip tends to equalize their forecast performance.

Although the choice of demand signal with consideration to bullwhip is important, forecast model can influence forecast error as well. Results from our regression analysis suggest that the additive model tends to increase forecast error. However, the additive model still outperformed multiplicative 44% of the time, often with significant improvement to forecast error. Our exploratory logistic regression analysis provides more insight on factors that contribute to selecting one model over the other.

Two key differences demarcate the two forecast models. Whereas the additive model's calculation weighs the *difference* between observed demand and forecasted components, the multiplicative model weighs the *ratio* of observed demand and forecasted components. Thus, the multiplicative model is more responsive to increased variance while the additive model is less so. In our logistic regression model, the only factor to negatively impact the likelihood of multiplicative model outperforming additive is trend. This is expected, as an increase in trend, all else equal, decreases the emphasis on variance, which compromises the value in using the multiplicative model. Particularly interesting is the positive association between POS and the preference for multiplicative model. This result seemingly suggests that POS's seasonal factor is more closely associated with multiplicative seasonal factor. In addition, bullwhip increases variance, for which the multiplicative model is well-equipped to address.

In consideration of both sets of regression results, we argue that while order data is often the preferred forecast input for suppliers attempting to forecast customer demand, POS's value in mitigating the destabilizing effect of bullwhip increases as demand variance becomes amplified. In addition, the preferred model is no longer multiplicative if POS were utilized. Overall, a two-step decision in the selection of demand signal and seasonal forecasting model can be formulated. First, the demand signal of choice should be decided based upon the degree of bullwhip. If the bullwhip effect is high, then POS can potentially result in more accurate forecast; otherwise order data should be utilized. The next decision would be the appropriate seasonal forecasting model. If POS data is utilized due to high degree of bullwhip, then the additive model should be selected. However, the multiplicative model is more likely to yield superior forecast performance in all other seasonal forecast settings.

H. Conclusions, Limitations, and Future Research

In this study, we attempt to examine the influence of demand signal and bullwhip on forecast accuracy and model choice for forecasting seasonal customer demand. While the value of downstream demand signal had been called into question, our results provide a clearer view to the picture. POS can remain an appropriate and effective demand signal when the bullwhip effect is significant. In addition, considering that the calculation of bullwhip requires visibility to downstream demand signals, information sharing in the retail industry has innate value even if POS is not utilized as a forecast basis.

Although results from this study are drawn from thirty-six distinct time series, they are all similar products competing in the same category, and sold through one retail format. Considering the diverse range of retail management policies that may impose influence on order variance, further studies are warranted with other seasonal categories, from other retailers and retail formats, with different demand signal both from the consumer level and the retail orders.

I. REFERENCES

Agarwal, A., & Holt, G. (2005), Reducing inventory by simplifying forecasting and using point of sale data. *MLOG Thesis*, Massachusetts Institute of Technology.

Allred, C.R., Fawcett, S.E., Wallin, C., & Magnan, G.M. (2012), A dynamic collaboration capability as a source of competitive advantage. *Decision Sciences*, 42, 129-161.

Autry, C.W., Griffis, S.E., Goldsby, T.J., & Bobbitt, L.M. (2005), Warehouse management systems: resource commitment, capabilities, and organizational performance. *Journal of Business Logistics*, 26 (2), 165-183.

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

Cachon, G., & Fisher, M. (2000), Supply chain inventory management and the value of shared information. *Management Science*, 46, 1032-1048.

Cao, M. & Zhang, Q. (2011), Supply chain collaboration: impact on collaborative advantage and firm performance. *Journal of Operations Management*, 29 (3), 163-180.

Chatfield, C. & Yar, M. (1988), Holt-Winters forecasting: some practical issues. *Journal of the Royal Statistical Society*, 37 (2), 129-140.

Chatfield, C. (1978), The Holt-Winters forecasting procedure. *Applied Statistics*, 27, 264-279.

Chen, L., & Lee, H.L. (2012), Bullwhip effect measurement and its implications. *Operations Research*, 60, 771-784.

Dejonckheere, J., S.M. Disney, M.R. Lambrecht, & Towill, D.R. (2003), Measuring and avoiding the bullwhip effect: A control theoretic approach. *European Journal of Operations Research*, 147 (3), 567-590.

Dekimpe, M.G., & Hanssens, D.M. (1995), The persistence of marketing effects on sales. *Marketing Science*, 14 (Winter), 1-21.

Dooley, K. J., T. Yan, S. Mohan, & Gopalakrishnan, M. (2010), 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.

Evers, P.T., & Beier, F.J. (1998), Operational aspects of inventory consolidation decision making. *Journal of Business Logistics*, 19, 173-189.

Germain, R., & Iyer, K.N.S. (2006), The interaction of internal and downstream integration and its association with performance. *Journal of Business Logistics*, 27 (2), 29-53.

Goldsby, T.J., Griffis, S.E., & Roath, A.S. (2006), Modeling lean, agile and leagile supply chain strategies. *Journal of Business Logistics*, 27, 57-80.

Greene, W.H. (2011), *Econometric Analysis*. 7th ed. Upper Saddle River, NJ: Prentice Hall.

Hanssens, D.M. (1998), Order forecasts, retail sales, and the marketing mix for consumer durables. *Journal of Forecasting*, 17, 327-346.

Kaipia, R., & Hartiala, H. (2006), Information-sharing in supply chains: five proposals on how to proceed. *International Journal of Logistics Management*, 17 (3), 377-393.

Kaipia., R., H. Korhonen, & Hartiala, H. (2006), Planning nervousness in a demand supply network: An empirical study. *International Journal of Logistics Management*, 17 (1), 95-113.

Lapide, L. (1998), New developments in business forecasting. *Journal of Business Forecasting Methods & Systems*, 17 (4), 29-30.

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

Lee, H.L., So, K.C., & Tang, C.S. (2000), The value of information sharing in a two-level supply chain. *Management Science*, 46 (5), 626-643.

McCarthy, T.M., Davis, D.F., Golicic, S.L., & Mentzer, J.T. (2006), The evolution of sales forecasting management: a 20-year longitudinal study of forecasting practices. *Journal of Forecasting*, 25 (5), 303-324.

McCullen, P., & Towill, D.R. (2002), Diagnosis and reduction of bullwhip in supply chains. *Supply Chain Management: An International Journal*, 7 (3), 164-179.

McKenzie, E. (1976), Error analysis for winters' additive seasonal forecasting system. *International Journal of Forecasting*, 2 (3), 373-382.

Mentzer, J.T., & Kahn, K.B. (1995), Forecasting technique familiarity, satisfaction, usage, and application. *Journal of Forecasting*, 14, 465-476.

Metters, R. (1997), Quantifying the bullwhip effect in supply chains. *Journal of Operations Management*, 15 (2), 89-100.

Olivia, R., & Watson, N. (2011), Cross-functional alignment in supply chain planning: a case study of sales and operations planning. *Journal of Operations Management*, 29, 434-448.

Rexhausen, D., Pibernik, R., & Kaiser, G. (2012), Customer-facing supply chain practices—the impact of demand and distribution management on supply chain success. *Journal of Operations Management*, 30, 269-281.

Schoenherr, T., & Swink, M. (2012), Revisiting the arcs of integration: cross-validations and extensions. *Journal of Operations Management*, 30, 99-115.

Silver, E.A., Pyke, D.F., & Peterson, R. (1998), *Inventory Management and Production Planning and Scheduling*, 3rd ed., New York: John Wiley & Sons.

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

Thornhill, N.F., & Naim, M.M. (2006), An exploratory study to identify rogue seasonality in a steel company's supply network using spectral principal component analysis. *European Journal of Operational Research*, 171 (1), 146-162.

Vogt, J.J. (2010), The successful cross-dock based supply chain. *Journal of Business Logistics*, 31, 99-119.

Williams, B.D., & Waller, M.A. (2010), Creating order forecasts: point-of-sale or order history? *Journal of Business Logistics*, 31, 231-251.

Williams, B.D., & Waller, M.A. (2011), Top-down vs. bottom-up demand forecasts: the value of shared point-of-sale data in the retail supply chain. *Journal of Business Logistics*, 32, 17-26.

Zinn, W., & Bowersox, D.J. (1988), Planning physical distribution with the principle of postponement. *Journal of Business Logistics*, 9, 117-136.

Table 1 – Descriptive statistics

	Mean	Median	Stand. Dev.	MAPE	Bullwhip	Var
MAPE	0.78	0.49	1.93			
Bullwhip	1.27	1.28	0.24	0.10		
VAR	189912.81	43826.29	342834.90	0.001	-0.30	
Trend	19.27	12.17	20.22	-0.08	-0.49	0.87

Table 2 – Differences in forecast performance

Method	Multiplicative		Additive	
	POS	Order	POS	Order
% Superior	62%	38%	63%	37%
% Difference	19%	22%	22%	43%
Demand Signal	POS		Order	
	Multiplicative	Additive	Multiplicative	Additive
% Superior	56%	44%	56%	44%
% Difference	75%	18%	62%	13%

Table 3 – Regression results for MAPE

DV = MAPE	Model 1	Model 2	Model 3	Alt. Model[a]
Control Variables				
Var	0.001 (0.000) *	0.001 (0.000) *	0.001 (0.000) *	0.001 (0.000) *
Trend	-0.024 (0.002) *	-0.029 (0.002) *	-0.035 (0.002) *	-0.049 (0.006) *
Horizon	0.094 (0.008) *	-0.039 (0.013) *	-0.054 (0.013) *	-0.056 (0.013) *
Alpha	1.181 (0.080) *	0.928 (0.081) *	0.898 (0.081) *	0.894 (0.081) *
Beta	1.070 (0.163) *	0.752 (0.164) *	0.712 (0.164) *	0.709 (0.164) *
Gamma	0.422 (0.037) *	0.376 (0.037) *	0.370 (0.037) *	0.369 (0.038) *
AR(1)	0.565 (0.005) *	0.552 (0.005) *	0.552 (0.005) *	0.544 (0.006) *
Effects of Interest				
Additive		0.345 (0.038) *	0.331 (0.038) *	0.329 (0.038) *
Bullwhip		0.429 (0.050) *	0.558 (0.055) *	0.054 (0.150)
POS		0.211 (0.049) *	1.437 (0.209) *	0.934 (0.248) *
Interaction				
*POS*Bullwhip*			-0.963 (0.160) *	-0.668 (0.188) *
Durbin-Watson	2.16	2.15	2.15	2.14
R-Squared	0.329	0.336	0.338	0.342
AIC	3.752	3.742	3.74	3.734

[a]specified with product fixed effects

*indicates variable significant at 0.01 level.

Table 4 – Exploratory logistic regression for model selection

DV = AMM	Model 1		Model 2	
Direct Effects				
C	-4.386 (0.091)	**	-3.573 (0.137)	**
Var	0.001 (0.000)	**	0.001 (0.000)	**
Trend	-0.102 (0.002)	**	-0.148 (0.004)	**
Horizon	0.214 (0.006)	**	0.209 (0.009)	**
Bullwhip	3.366 (0.061)	**	0.314 (0.091)	**
POS	0.360 (0.022)	**	0.479 (0.193)	*
Interaction Effects				
*POS*Bullwhip*			-0.251 (0.126)	*
*POS*Trend*			-0.050 (0.006)	**
*POS*Var*			0.001 (0.000)	**
*POS*Horizon*			0.054 (0.014)	**
McFadden R-Squared	0.425		0.475	
-Log Likelihood	9291.083		8486.294	
AIC	0.797		0.728	

**p<0.01; *p<0.05

Figure 1 – DF-competition design

	Additive	Multiplicative
POS	Input Data: POS Model: Additive	Input Data: POS Model: Multiplicative
Order	Input Data: Order Model: Additive	Input Data: Order Model: Multiplicative

Figure 2 – Average MAPE comparison by forecast method

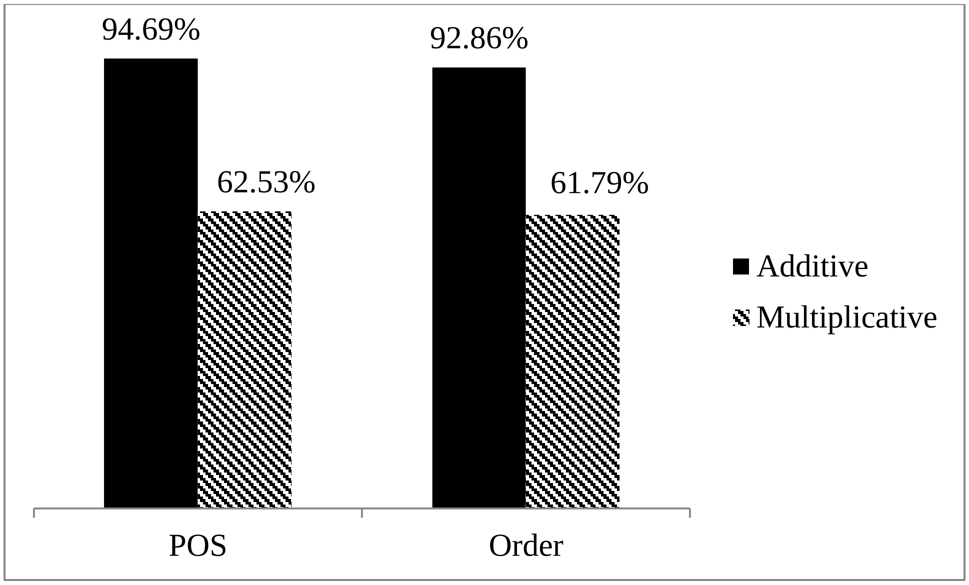
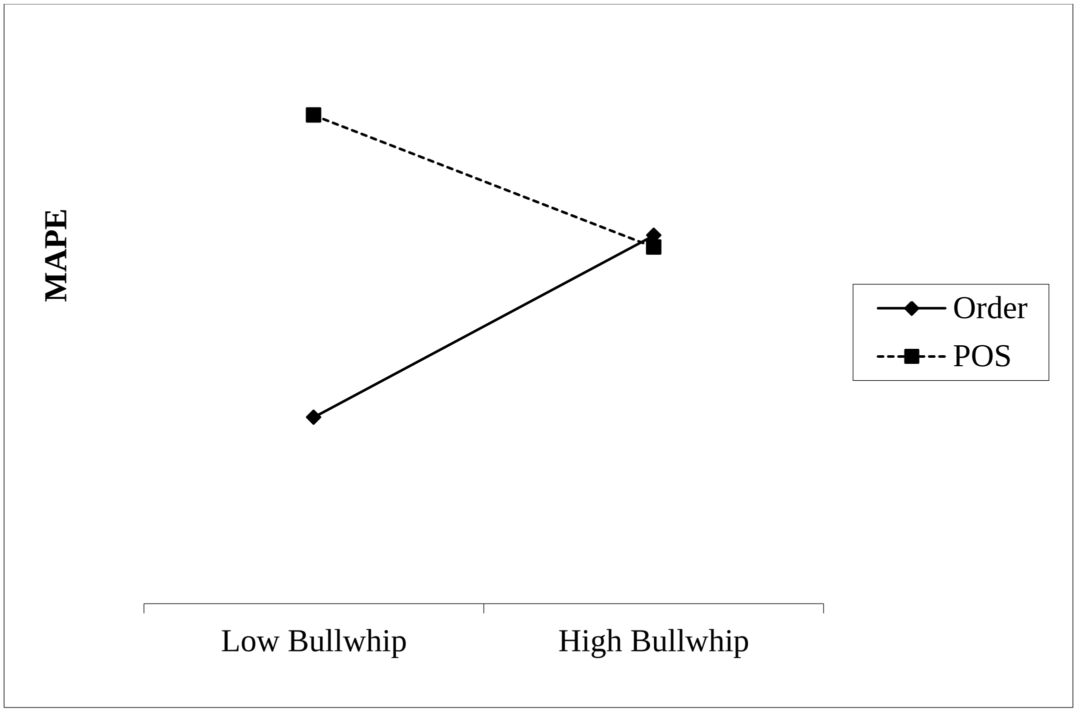


Figure 3 – Two-way interaction plot for demand signal and bullwhip



Appendix I – DF-competition design

Parameter Name	# of Parameters	Parameters
Forecast Models	2	*Additive, Multiplicative*
Demand Signals	2	*POS, Order*
Alpha Levels	3	*0.51, 0.19, 0.02*
Beta Levels	3	*0.176, 0.053, 0.005*
Gamma Levels	3	*0.5, 0.1, 0.05*
DCs	6	
Products	6	
Out of Sample Forecasts	6	
Total Observations (N)	23,328	

Chapter 6 – CONCLUDING DISCUSSION

VI. CONCLUDING DISCUSSION

With the rise of "the internet of things" (Chui et al., 2010), data-driven decisions are expected to become an even more important way through which companies may obtain and sustain competitive advantage over their rivals. Already, companies such as Wal-Mart and Amazon.com have reaped tremendous benefit from possessing a greater understanding of how to leverage data to formulate and execute supply chain decisions (Bollier, 2010). In turn, many suppliers in the CPG industry also sought to obtain similar benefits through collaborative supply chain management with their key retail customers (Hofer et al., 2012). Through expensive IT investments (Ravichandran & Liu, 2011), suppliers are able to minimize waste and raise their service levels to generate greater revenue without proportionate increase in cost (Baker, 2008). Today, suppliers such as ConAgra Foods and Coca Cola continue to actively explore ways to further utilize supply chain data for increased effectiveness in demand planning. Achieving such goals require companies to step away from a "black box" approach and instead become more methodical in supply chain data science.

This dissertation examined various countervailing statistical effects that may confound supply chain performance metrics as well as demand planning. Considering the degree of importance placed by most companies on these statistical outcomes (Rexhausen et al., 2012), the potential negative impact from misguided actions can be particularly damaging to the supply chain. Therefore, a primary goal of this dissertation was to examine how these statistical effects influence a key measure of supply chain volatility, namely the bullwhip effect, and customer demand forecasting. Accurate measurement of volatility allows companies to better gauge the value of various sources of information and formulate superior capacity planning. In addition, this dissertation also explored the potential for these statistical effects to be leveraged as tools for planners to achieve superior demand forecasts, which enables companies to more efficiently position inventory throughout their distribution networks.

In particular, three topics were examined by the three essays in this dissertation. Essay 1 examined data aggregation as an explanation for the conflicting empirical literature on the prevalence of the bullwhip effect. Next, essay 2 explored two countervailing statistical effects of temporal aggregation on forecast accuracy, as well as their moderating effects on the relationship between demand signal and forecast accuracy. Finally, essay 3 first challenged the conventional notion that downstream demand signal is always superior as a forecast input, and then examined factors that determine selecting between two seasonal forecasting models. A summary of each essay will be discussed below.

In essay 1 (Figure 1), it was noted that Cachon et al. (2007) and Bray and Mendelson (2012) arrived at conflicting conclusions regarding the magnitude and prevalence of the bullwhip effect at the industry and firm levels, respectively. In response, Chen and Lee (2012) analytically demonstrated how both product-location and temporal aggregation may mask the bullwhip effect. Utilizing a large set of order and POS data for three categories of products, collected from regional DCs operated by a large national retailer, essay 1 first empirically validated Chen and Lee's (2012) propositions on the effects of data aggregation. Furthermore, essay 1 also corroborated with Bray and Mendelson (2012), which stated that seasonality is an important determinant of the bullwhip effect as well. Overall, essay 1 showed that the conflicting observations made by Cachon et al. (2007) and Bray and Mendelson (2012) is primarily due to their different levels of analysis and the degree of seasonality within product categories. Finally, essay 1 noted that fundamental differences exist among the three widely-accepted measures of bullwhip.

Essay 2 (Figure 2) further examined temporal aggregation's effect in the context of forecasting. In particular, two effects of temporal aggregation—information loss and variance reduction—were hypothesized to affect forecast accuracy. Whereas the information loss effect (e.g., Amemiya & Wu, 1972; Rossana & Seater, 1995) argues that temporal aggregation results in altered statistical properties to increase forecast error, the variance reduction effect (Hotta et la., 2005) posits that temporal aggregation enhances data stability to decrease forecast error. The two countervailing arguments were tested through a quasi-experiment by generating customer demand forecast with both order and POS data for products from two non-seasonal categories. Results suggest that depending on the demand signal and its degree of autocorrelation, either statistical effect can dominate the other to determine the temporal aggregation's overall effect on forecast error. Specifically, while the information loss effect is dominant when POS data is utilized, it is eclipsed by the variance reduction effect when order data is used. Furthermore, the variance reduction effect is amplified as autocorrelation becomes increasingly perfectly negative.

Essay 3 built on essay 2 in two ways (Figure 3). First, the conventional wisdom, that downstream demand signals are generally preferred forecast input (e.g., Lee et al., 1997), is tested in seasonal forecasting. Results showed that forecasting customer demand for seasonal products should generally utilize order, rather than POS data. However, forecast performance gap between order and POS data diminishes as bullwhip increases. Second, essay 3 also examined the factors that may help forecasters to choose between additive and multiplicative forecasting models. Results suggested that while the multiplicative model is generally preferred for downstream demand signals, this relationship is influenced by bullwhip. Increased demand distortion destabilizes forecast, therefore the additive model, which tends to yield a more conservative estimate of the seasonal factor, may be preferred over the multiplicative model for downstream demand signals.

Overall (Figure 4), the dissertation showed that as supply chain management becomes increasingly driven by data, countervailing statistical effects can impact both demand distortion metrics and demand forecasts. Therefore, a scientific approach to utilizing supply chain data is necessary for performance gains. This dissertation identified and reconciled three key statistical effects to reach the following conclusions. First, the level of analysis can have substantial influence on the observance of demand distortion in the supply chain. Findings corroborate with both Zotteri and Kalchschmidt (2009) and Chen and Lee (2012) to reinforce the importance of alignment between the level of analysis and the level of decision. In addition, temporal aggregation is a double-edged sword in forecasting. While temporal aggregation can benefit demand forecasting by reducing data volume and stochastic variance, it can also have the opposite effect due to information loss. Although both statistical effects are concurrent, the overall impact on statistical forecast accuracy is determined by a combination of demand signal selection and its autocorrelation factor. And lastly, while POS may not be the best forecast input for forecasting seasonal retail orders, the advantage of order data becomes increasingly dubious as bullwhip increases. This effect can impact both forecast model choice and forecast accuracy.

A. THEORETICAL CONTRIBUTIONS

Companies invest a tremendous amount of resources in hopes of planning and managing a superior supply chain. Yet, empirical and anecdotal evidence show that supply chain integration and collaboration are both difficult to establish and even harder to translate to expected performance gains (Fawcett & Magnan, 2002; Jin et al., 2013). Enabled by various industry initiatives and alliances between supply chain partners, collaborative demand management strategies rapidly increased the upstream flow of downstream demand signals (e.g., Waller et al., 1999; Frankel et al., 2002). If properly utilized, theses data can allow firms to anticipate and mitigate market uncertainties (Ravichandran & Liu, 2011; Rexhausen et al., 2012), improve dynamic collaboration capabilities (Allred et al., 2012), and forge more enduring and fruitful supply chain integration efforts (Schoenherr & Swink, 2012). Furthermore, when relevant data is systematically shared and collaboratively utilized, benefits such as reduced purchasing, inventory and distribution costs (Williamson et al., 1990; Baker, 2008) can increase firm performance (Bower, 2006; Muzumdar & Fontanella, 2006). While much of the current literature on supply chain management emphasizes how data should be shared (e.g., Tohamy, 2008; Atkinson, 2009), only a few studies exist on supply chain data science—how shared data should be scientifically utilized (e.g., Williams & Waller, 2010; 2011).

This dissertation makes several contributions to the theoretical literature on supply chain management, specifically in the burgeoning literature on supply chain data science (e.g., Williams & Waller, 2010; 2011). First, it was found that conflicting empirical observations of bullwhip is explained by both product-location and temporal aggregation. Reconciling this conflict revealed insight into how statistical aggregation may influence measures of supply chain volatility. Particularly pertinent to supply chain management is that bullwhip has significant cost implications (Lee et al., 2000), and a first step toward mitigating bullwhip is accurate measurement. Therefore, the lack of alignment between the level of measurement and the level decision will lead to suboptimal decisions regarding capacity and supply. Results from this study complement the existing supply chain literature on relational and process integration by demonstrating that statistical influences may either lead or hinder supply chain performance.

According to Croxton et al. (2002), successful demand planning in the supply chain requires accurate forecasting to synchronize supply and demand. When substantial bullwhip exists in the supply chain, "grossly inaccurate demand forecasts" can lead to "low capacity utilization, excessive inventory, and poor customer service" (Lee et al., 2000, p. 626). Although the use of downstream demand signals can help suppliers to mitigate bullwhip's negative influence (Lee et al., 1997; Cachon & Fisher, 2000), results from this dissertation suggest that selective use of temporal aggregation may be a viable alternative. Specifically, while temporal aggregation can reduce variance that was amplified by demand distortion, this benefit may not exceed the harm caused by information loss.

Lastly, one of the less explored aspects of demand planning is seasonal forecasting. Although downstream demand signals are generally believed to be a superior source of information for forecasting customer demand (Cachon & Fisher, 2000; Lee et al., 2000), their advantage may not hold true when seasonality exists (Williams & Waller, 2010). Results from this dissertation suggest that seasonal forecast accuracy depends on a complex mix of factors including the choice of demand signal, the degree of demand distortion, both of which determine the optimal forecast model. For seasonal forecasting, while customer order data tend to outperform downstream demand signal, their difference diminishes as demand distortion increases. In addition, while the multiplicative forecast model generally outperforms the additive, if downstream demand signal is favored due to demand distortion, then the additive model is preferred over the multiplicative.

B. MANAGERIAL IMPLICATIONS

Several findings from this dissertation may assist firms in applying data science to supply chain demand management. Obtaining the necessary resources and capabilities to access and utilize downstream demand signals can be costly for both suppliers and retailers. Therefore, a methodical approach to analyzing supply chain data allows companies to maximize the value of shared information. To begin, the level at which to measure demand distortion should be considered in conjunction with the level of the decision. Although aggregation results in fewer data points to reduce computational and resource intensity, it may result in underestimating demand distortion.

For the product-location level of data, analysis should be conducted for the level at which replenishment occurs. For example, if a supplier is replenishing two different DCs, estimating demand distortion using their aggregated data would result in masking the underlying demand volatilities at each location individually. In other words, statistically aggregating data series from both locations in effect treats two points of demand as a single consolidated location, at which point risk-pooling occurs (e.g., Zinn et al., 1989) to mask true underlying volatility (Chen & Lee, 2012). From a demand planning standpoint, the supplier might be misled into underestimating both the capacity necessary for achieving desired service levels at both locations as well as the potential improvement to demand forecast using POS data.

Temporal aggregation can also mislead suppliers as outlined above. But from a forecasting perspective, it may also be selectively used as a tool for mitigating bullwhip. As demand signals are processed and formulated into orders, managerial and behavioral influences can induce variance to mislead future forecast. Many smaller suppliers and retailers lack the resources and capability to share information and engage in collaborative replenishment and distribution activities. In lieu of such resource intensive strategies, temporal aggregation may be used to minimize demand distortion's destabilizing effect on future forecasts. This statistical method can be especially beneficial to those data series that are highly negatively autocorrelated. This benefit may also extend to downstream demand signals, but only when the autocorrelation factor of the data series is highly negative. That is because downstream demand signal's principle value is its reflection of consumer demand (e.g., Lee et al., 2000). Therefore the net impact of temporal aggregation, when forecast input is the downstream demand signal, can be positive only when the detrimental impact of information loss is offset by the benefit of variance reduction. Moreover, results also show a common misconception among theorists and practitioners. Contrary to the belief that downstream demand signal is always superior, both temporally aggregated and disaggregated order data outperform disaggregated POS data when it is highly negatively autocorrelated.

Taking customer demand forecasting to a seasonal product setting, our results once again show that downstream demand signal is not always the superior forecast input. Demand planners should be cognizant of whether the retail customer's ordering policies are relatively stable over time. If so, then historic order data would likely show recurring patterns that reflect the "rhythms" established by previous retail seasonal smoothing processes. These patterns can be utilized by planners to forecast future orders. Furthermore, multiplicative model is more responsive to changes to seasonality, which makes it ideal for short-term seasonal demand forecasts. However, if the retail customer's policies are relatively idiosyncratic, then stochastic variance amplification results in heightened bullwhip effect. To obtain a more conservative estimate of the seasonal demand and avoid overreaction to idiosyncratic fluctuations, planners may utilize POS data and the additive model.

Finally, as outlined in Figure 4, the process of translating business data to actionable intelligence is highly complex and requires a methodical approach. In general, supply chain partners would be prudent to share data. Retailers can empower suppliers by sharing downstream demand data. Correct levels of analysis enable accurate assessment of both demand distortion and demand characteristics, which in turn determine the proper forecast model. Finally, all of these factors—level of analysis, demand characteristics, and forecast model—collectively assist suppliers in stepping away from the "black box" approach and leverage supply chain data science to plan for future retail demand.

C. REFERENCES

Allred, C.R., Fawcett, S.E., Wallin, C., & Magnan, G.M. (2012). A dynamic collaboration capability as a source of competitive advantage. *Decision Sciences*, 42, 129-161.

Amemiya, T., & Wu., R.Y. (1975). The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association*, 67 (39), 628-632.

Atkinson, W. (2009). S&OP: now more than ever. *Supply Chain Management Review*, 13, 50-54.

Baker, P. (2008). The design and operation of distribution centers within agile supply chains. *International Journal of Production Economics*, 111, 27-41.

Bollier, D. (2010) *The Promise and Peril of Big Data*. The Aspen Institute, Washington DC.

Bower, P. (2006). How the S&OP process creates value in the supply chain. *Journal of Business Forecasting*, 25, 20-32.

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

Cachon, G., & Fisher, M. (2000). Supply chain inventory management and the value of shared information. *Management Science*, 46, 1032-1048.

Cachon, G. P., Randall, T., & Schmidt. G.M. (2007). In search of the bullwhip effect. *Manufacturing and Service Operations Management*, 9 (4), 457-479.

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

Chui, M., Loffler, M., & Roberts, R. (2010). The Internet of Things. *McKinsey Quarterly*. March, 1-14.

Croxton, K.L., Lambert, D.M., Garcia-Dastugue, S.J., & Rogers, D.S. (2002). The demand management process. *International Journal of Logistics Management*, 13, 51-66.

Fawcett, S.E., & Magnan, G.M. (2002). The rhetoric and reality of supply chain integration. *International Journal of Physical Distribution and Logistics Management*, 32 (5), 339-361.

Frankel, R., Goldsby, T.J., & Whipple, J.M. (2002). Grocery industry collaboration in the wake of ECR. *International Journal of Logistics Management*, 13 (1), 57-71.

Hofer, C., Jin, H., Swanson, D., Waller, M.A., & Williams, B. (2012). The impact of key retail accounts on supplier performance: a collaborative perspective of resource dependency theory. *Journal of Retailing*, 88, 412-420.

Hotta, L.K., Valls Pereira, P.L., & Otta, R. (2005). Effect of outliers on forecasting temporally aggregated flow variables. *Sociedad de Estadistica e Investigacion Operativa Test*, 13, 371-402.

Jin, Y.H., Fawcett, A.M., & Fawcett, S.E. (2013), Awareness is not enough: commitment and performance implications of supply chain integration. *International Journal of Physical Distribution and Logistics Management*, 43 (3), 205-230.

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

Lee, H.L., So, K.C., & Tang, C.S. (2000). The value of information sharing in a two-level supply chain. *Management Science*, 46 (5). 626-643.

Muzumdar, M., & Fontanella, J. (2006). The secrets to S&OP success. *Supply Chain Management Review*, 10, 34-41

Ravichandran, T., & Liu, Y. (2011). Environmental factors, managerial processes, and information technology investment strategies. *Decisions Sciences,* 42, 537-574.

Rexhausen, D., Pibernik, R., & Kaiser, G. (2012). Customer-facing supply chain practices – the impact of demand and distribution management on supply chain success. *Journal of Operations Management*, 30, 269-281.

Rossana, R.J., & Seater, J.J. (1995). Temporal aggregation and economic time series. *Journal of Business & Economic Statistics,* 13, 441-451.

Schoenherr, T. & Swink, M. (2012). Revisiting the arcs of integration: cross-validations and extensions. *Journal of Operations Management,* 30, 99-115.

Tohamy, N. (2008). The evolution of S&OP. *Supply Chain Management Review*, 12, 10-11.

Waller, M.A., Johnson, E.M., & Davis, T. (1999). Vendor-managed inventory in the retail supply chain. *Journal of Business Logistics*, 20 (1), 183-203.

Williams, B.D., & Waller, M.A. (2010). Creating order forecasts: point-of-sale or order history? *Journal of Business Logistics,* 31, 231-251.

Williams, B.D., & Waller, M.A. (2011). Top-down vs. bottom-up demand forecasts: the value of shared point-of-sale data in the retail supply chain. *Journal of Business Logistics,* 32, 17-26.

Williamson, K.C., Spitzer Jr., D.M., & Bloomberg, D.J. (1990). Modern logistics systems: theory and practices. *Journal of Business Logistics*, 11, 65-87.

Zotteri, G., & Kalchschmidt, M. (2007). Forecasting practices: empirical evidence and a framework for research. *International Journal of Production Economics,* 108, 84-99.

Figure 1 – Essay 1 diagram

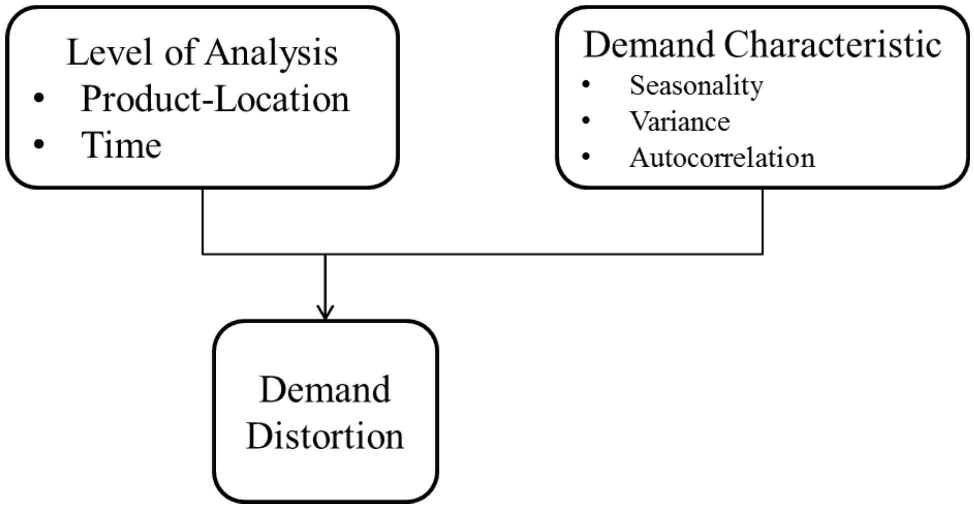


Figure 2 – Essay 2 diagram

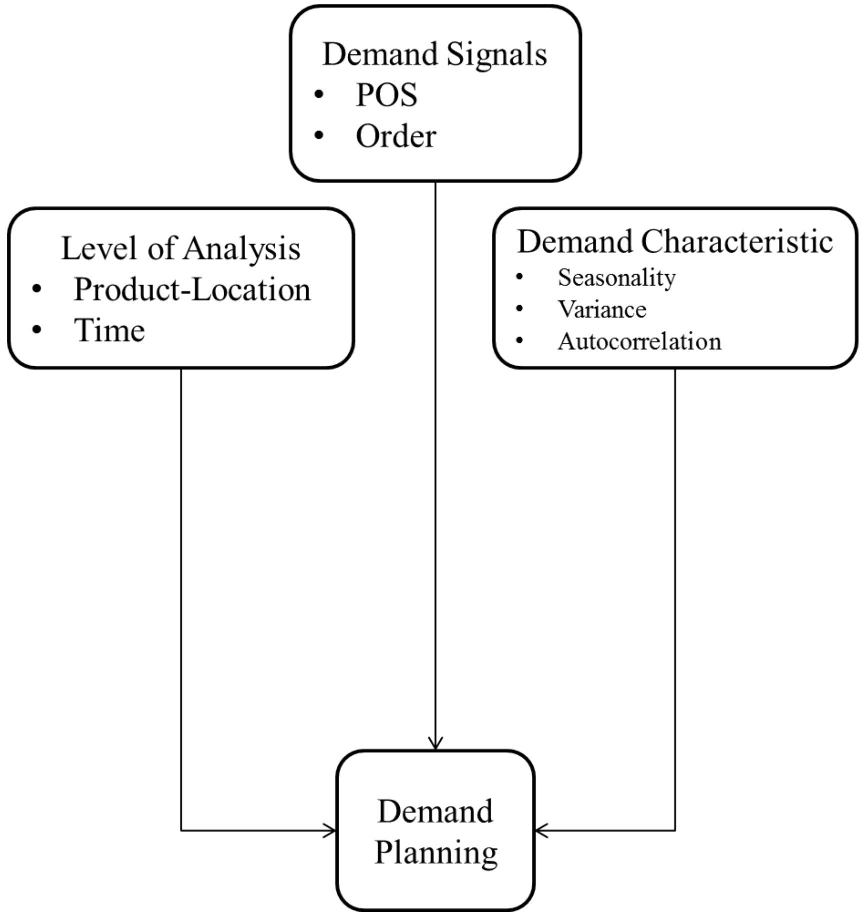


Figure 3 – Essay 3 diagram

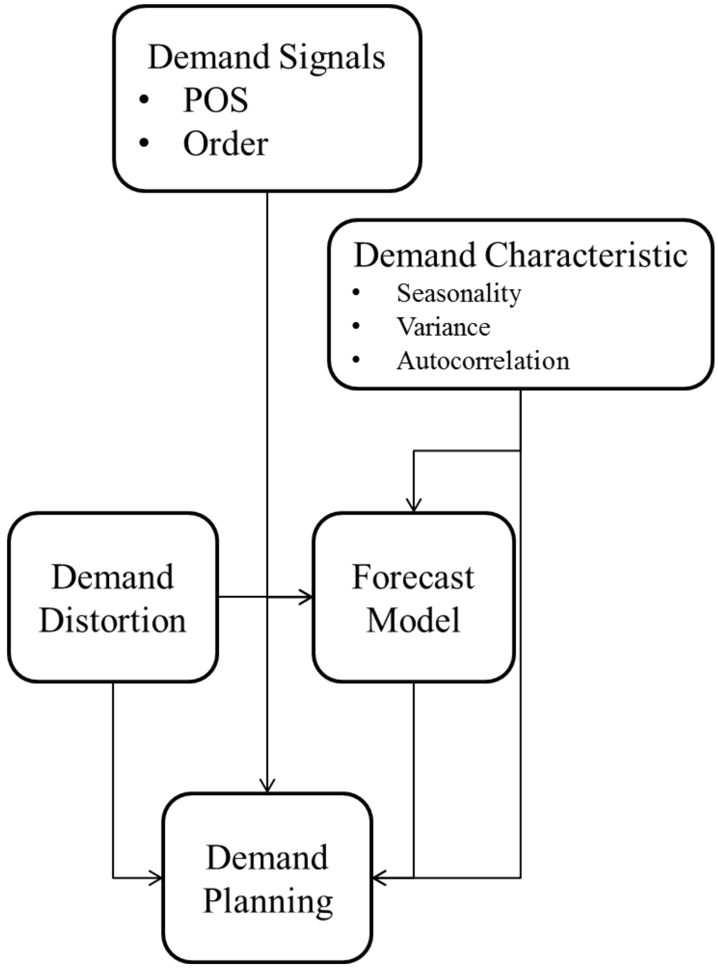
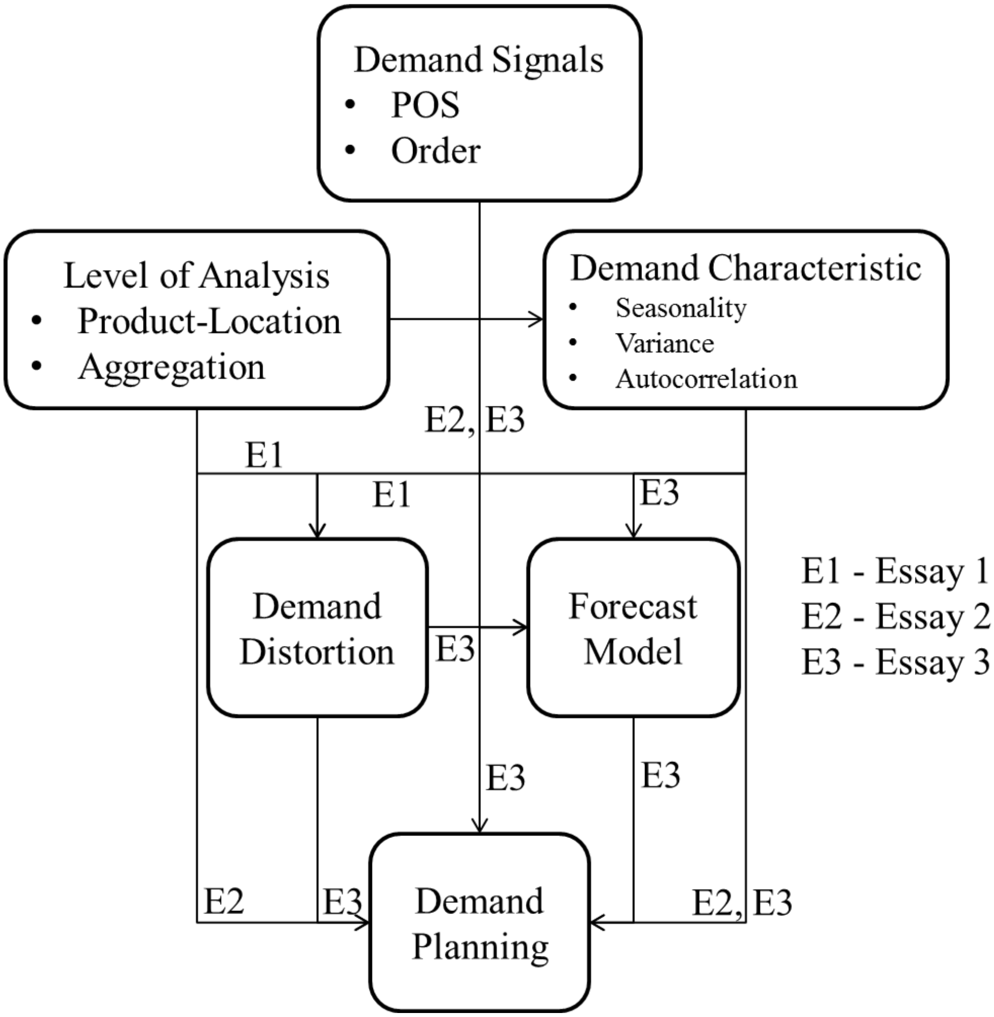


Figure 4 – Overall dissertation diagram



Chapter 7 – FUTURE RESEARCH

VII. FUTURE RESEARCH

A. DATA SCIENCE 1 – LOSS INTEGRAL

The concluding figures in this dissertation point to potentially greater streams of research in supply chain data science. Throughout the dissertation, data aggregation took a central role as a primary influence on various metrics of supply chain managements. While improving metrics such as demand distortion and forecast error through proper specifications of data and model parameters can result in increased service levels, stock-outs may still occur. Each time when stock-out occurs, several negative outcomes may have ripple effects to further distort demand in the supply chain.

First, substitute purchases inflate expected future demand for that product. As a result, demand variance for the substitute product also increases. Since retailers set inventory based on expected demand, greater variance will result in lower service levels without greater inventory. If stock-out occurs for the substitute product, a contagion effect may result to spread to other substitute products as well. This is a potentially serious consequence, since demand variance for other products may increase to result in greater supply chain costs. However, if the retailer attempts to anticipate this effect by increasing inventory for the substitute product in advance, overstock may potentially occur. The major costs incurred are thus inventory costs at the retailer level and inflated bullwhip to the suppliers. If no structural changes occur to the demand series for the substitute product, then it can be argued that several periods later, this problem will correct itself as demand and ordering both return to their normal state.

Second, lost sales hinder visibility to true underlying demand. If the consumer chooses to either purchase a substitute product or skip this purchase cycle altogether, then observed demand for the focal product is likely not sufficient to generate accurate forecast for the next period. In turn, demand planners are likely to generate inadequate levels of future demand forecast, thereby increasing the likelihood of future stock-outs. The consumer would then likely to purchase products elsewhere, leading to entirely new shopping habits and cause the retailer to lose out not only on a single item worth of sales, but an entire basket as well. Alternatively, if the consumer continues to purchase substitute products, then the first consequence would likely to become even worse.

The third point has interesting implications. While POS data is believed to be largely free of the bullwhip effect, stock-outs may delay purchases to force cyclicality and other data distribution properties that would not otherwise exist. For example, for many retailers that rely on price promotions, sales tend to track retail operations in two ways. First, retailers such as drugstores receive weekly or bi-weekly replenishment from their warehouses. As a result, their sales for many products tend to follow this schedule: stock-out occurs one or two days before replenishment, causing customers to delay their purchases. In addition, price promoters use loss leaders to draw store traffic. Hence, their sales tend to follow their promotions as well. Further complicating the matter is the practice of giving “rain checks” to consumers. When a stock-out occurs for promoted items, many hi-lo retailers give coupons to consumers to honor their sale price in the future when replenishment stock arrives. This is problematic in several ways. First, it induces bullwhip even at the store level. Second, many suppliers give price breaks to the retailer for such promotions. Replenishment stock is not likely purchased under price breaks, thereby causing the retailer to incur a loss when consumers return to purchase the product at the sale price. Third, because most hi-lo retailers repeat annual sales events on similar products, their systems would continue to order at quantities below the true expected demand since they do not account for rain check sales past the promotional period. Therefore, of vital importance to the retail supply chain is the capability for a retailer to measure lost sales.

Loss Integral Background

The loss integral approximates the expected loss based on any given continuous distribution. It is a widely applied function across many fields, such as finance, learning, forecasting, and policy making. In supply chain management, the loss integral has been commonly used to investigate optimal inventory levels given various conditions and assumptions (e.g., Nahmias & Smith, 1994; Huh et al., 2009).

The principle function of inventory is to satisfy fluctuating demands. Retailers attempt to set inventory levels they deem appropriate for desired in-stock probabilities. In order to calculate in-stock probabilities, retailers must first forecast estimated demand based on time-series models of historical sales data. The vast majority of retailers forecast demand with such assumptions as spatial independence using simplistic statistical techniques based on normally distributed historical sales. Yet, historical retail sales are observed demand that is subject to truncation and censoring when stock-out occurs (Conrad, 1976). Thus, inventory policies based purely on censored demand is likely suboptimal and result in continued underestimation of true demand.

We model a retail store selling a single SKU where demand is continuous, stationary and nonnegative. The store uses (r, Q) continuous review model to replenish the SKU

Assumptions:

1. A continuous review system is used for replenishment with reorder point r
2. Demand is stochastic and stationary

3. Demand during lead time is a continuous random variable, X, that is nonnegative
4. The probability density function is given by $f(x)$, the cumulative distribution function is given by $F(x)$, and the expected value is given by $E(X)$

Proposition. The loss integral can be written as

$$L(r) = \int_{x=r}^{\infty} (x - r)\, f(x) dx$$

$$= E(X) - \int_{x=0}^{r} x\, f(x) dx - r\big(1 - F(r)\big).$$

Proof.

$$L(r) = \int_{x=r}^{\infty} (x - r)\, f(x) dx$$

$$= \int_{x=r}^{\infty} x\, f(x) dx - \int_{x=r}^{\infty} r\, f(x) dx.$$

Note that $E(X) = \int_{x=0}^{\infty} x\, f(x) dx = \int_{x=0}^{r} x\, f(x) dx + \int_{x=r}^{\infty} x\, f(x) dx$. This can be rewritten as $\int_{x=r}^{\infty} x\, f(x) dx = \int_{x=0}^{\infty} x\, f(x) dx - \int_{x=0}^{r} x\, f(x) dx$. Substituting this expression for $\int_{x=r}^{\infty} x\, f(x) dx$ in the above equation we have

$$L(r) = \int_{x=0}^{\infty} x\, f(x) dx - \int_{x=0}^{r} x\, f(x) dx - \int_{x=r}^{\infty} r\, f(x) dx$$

$$= \int_{x=0}^{\infty} x\, f(x) dx - \int_{x=0}^{r} x\, f(x) dx - r \int_{x=r}^{\infty} f(x)\, dx$$

$$= E(X) - \int_{x=0}^{r} x\, f(x) dx - r\left(1 - \int_{x=0}^{r} f(x)\, dx\right)$$

$$= E(X) - \int_{x=0}^{r} x\,f(x)dx - r\big(1 - F(r)\big).\,Q.E.D.$$

$L(r) = E(X) - \int_{x=0}^{r} x\,f(x)dx - r\big(1 - F(r)\big)$ can be used to estimate the expected number of units out from a historical perspective. Here is the process:

Step 1. Use the forecast of demand during the lead time as an estimate of $E(X)$.

Step 2. Calculate the average number of units sold during the previous lead times as an estimate of $\int_{x=0}^{r} x\,f(x)dx$.

Step 3. Calculate the percentage of times no stockouts occurred during the lead times as an estimate of $F(r)$.

Since r is known, use the three estimates above to estimate $L(r)$, by taking the forecast of demand during lead time in Step 1 and subtracting the estimate of average units sold during the lead time from Step 2. Then multiply the reorder point by the frequency of stockouts.

B. DATA SCIENCE 2 – CUSTOMER DEMAND FORECAST FOR PRODUCT LINE EXTENSIONS

In many disciplines, strategic management requires firms to assess both internal and external forces for decisions. A major component of strategic management in retail is segmentation, in which firms attempt to expand their consumer base. Often, new products are introduced with no prior sales history. As a result, demand planners have little to no guidance with regard to potential demand for the new product. Moreover, most demand forecast techniques are variants of simplistic exponential smoothing processes. Without sales history, such quantitative models require alternate data to approximate anticipated demand. In practice, demand planners frequently use sales history of a similar product, along with some component of qualitative reasoning, to form an estimated demand for the new product.

Segmentation may be achieved in several ways. A firm may adopt a completely new entity, or brand, to be marketed toward their targeted segment. One such example is Gap's high end Banana Republic stores. On the other hand, a firm may wish to leverage its existing core brand value in order to have a more evolutionary approach in its extension. The counterpart to Gap's Banana Republic may be Donna Karan's DKNY extension. However, regardless of the type of extension, marketers continue to grapple with potential pitfalls of inadequately planned extensions. One such result is cannibalization, when an extension usurps market share away from the incumbent brand.

However, firms rarely expand in to a completely foreign segment. Thus, segmentation strategies are implemented based on some original brand or product line that is already being marketed. In the segmentation process, inevitably some characteristic is carried over (Moorthy, 1984). As a result, it would be reasonable to anticipate the sales of specific segment to at least somewhat resemble demand patterns of the established product line or brand. Yet in the process of planning the distribution of a new extension, forecast basis for these products are often made by relying on a combination of arbitrary decisions, educated beliefs, and historical precedence of established brands. Evaluating the effectiveness of such a forecast approach is often done several periods after product launch. However, the uncertainty surrounding a new product's initial sales period command significant costs that are avoidable with improved demand planning. Therefore, it is important to explore how to incorporate past demand patterns from an established brand or product line through more sophisticated statistical techniques.

<u>Product Lines and Extensions</u>

A product line can be define as a group of products that are closely related, marketed through similar channels, fall within similar price ranges or sold to similar customer groups (Armstrong

& Kotler, 2006). Within these lines, products can be differentiated either vertically or horizontally (Randall et al., 1998). Vertical differentiation refers to variation within quality levels of products and horizontal differentiation refers to variation with the function or category of the product; this can also be referred to as quality-based segmentation (Desai, 2001). Hilton Worldwide is an expansive example of vertical differentiation in the hotel industry; their brands (e.g., Waldorft Astoria, Hilton, Embassy Suites, Hilton Garden Inn, Hampton, & Double Tree) intentionally differ in perceived quality. Although each hotel has the same function, they all differ in eminence.

Horizontal differentiation refers to variation with certain product characteristics to appeal to different target markets (Randall et al., 1998). An example of horizontal differentiation is Dove's deodorant product line that not only has several different types of deodorants (i.e., solids, aerosols, roll-ons & body mists), but each type of deodorant comes in different scents (e.g., original clean, fresh burst, wild rose, smooth cashmere); these products are all positioned as being of equal quality, and differ in terms of packaging, formulations, and applications. Thus, a horizontally-differentiated product tends to have a target segment of consumers that will otherwise not purchase the original product.

A product line extension occurs when a company adds more brands or models to its current product line (Solomon et al., 2009). According to Meyer et al. (1997) there are two criteria for a product to be considered a product line extension. First, the novel product must embody the core features on the already existing products. That is, the extension must be related to its predecessors. Second, the new product must target new customer segments than the existing products. When new products take away market share of existing products, cannibalization occurs; that is, competition among a firm's product line (Moorthy, 1984). These thoughts all resonate from the theory of segmentation.

Under the theory of segmentation, the overall market (e.g., the ready-to-eat cereal market) is viewed as heterogeneous, with several homogenous market segments or "wedge shaped pieces" within (Smith, 1956 pg. 5). Marketers then, are encouraged to create and position different products for each important market segment (Smith, 1956). Promotion is used heavily to inform each segment of the products that have been created to specifically meet their needs or wants (Smith, 1956). Ideally, products will not cannibalize each other because each product targets a different segment, and each segment will not be interested in other segments' products (Frank et al., 1972).

Although firms would enjoy perfect segmentation, it has been shown to be unrealistic (Moorthy, 1984). Often, horizontally-differentiated products do not perfectly establish a separate segment, which results in cross-segment consumption, i.e. product cannibalization. Due to the associated negative consequences (Solomon et al., 2009), cannibalization thus carries a heavily negative connotation and is simultaneously viewed as unavoidable (Moorthy, 1984). Therefore, while demand for the established product affect the demand for its extension due to shared product characteristics, the reverse is true as well. That is, due to the cannibalization effect, demand for the established product is also affected by the demand for its extension.

<u>Information Content of Product Variant Demand Signals</u>

Every data series for a demand signal contains information pertaining to customer demand as reflected by data variance. A long line of literature in demand forecasting supports the notion that demand signals observed at point-of-sale (POS) and distribution center (DC) levels each contain unique information (Williams and Waller, 2010; 2011). At the POS level, demand signals reflect consumer demand patterns such as paycheck cycles and seasonality. At the DC level, demand signals instead reflect warehouse management idiosyncrasies as well as other supply chain processes.

Forecasting demand for new product extensions is especially difficult. First, new products tend to have less demand due to a lack of market penetration. As a result, low sales volumes can potentially take the form of intermittent demand signals, which are notoriously difficult to forecast. Second, new products tend to have trend components to their demand characteristic that may not be present in demand signals for other more established products. As a result, demand planners may only formulate their best guess at potential trend.

Building on the product line extension line of literature, we argue that demand signals for established products at the POS level contains consumer purchase patterns specific to that product. Factors such as shelf-life, package size and purpose all determine the frequency and quantity at which consumers purchase the established product. Since a product line extension embodies the core features of the original product as well as share substantial similarities (Meyer et al., 1997), similar consumer purchase habits may be anticipated. Therefore, the demand signal for the established product can assist in forecasting sales of its extension.

Many companies tend to extend product lines that already have multiple iterations of the same base product. For example, Proctor and Gamble uses its Crest toothpaste as the established product for many extensions in mouthwash products instead of its lesser-known Scope brand. However, not all products in the same line have identical demand patterns. Since segmentation occurs in “wedges” (Smith, 1956), new product extensions likely appeal to consumers of the more established products as well. Therefore, demand signals from multiple products within the same product line can be beneficial for forecasting demand for new product extensions.

Information Content of Downstream Demand Signal

Downstream demand signals can be leveraged for superior customer demand forecast (Williams & Waller, 2010; 2011). More specifically, for non-seasonal product demand, POS data tends to outperform order data because bullwhip amplifies order data variance to destabilize order forecast. Principle statistical effect of bullwhip is amplified demand variance (Lee et al., 1997), which distorts two main components of demand signals—stochastic and seasonal variance (Metters, 1997; Bray & Mendelson, 2012). For both variance components, retail ordering policies may induce additional seasonality effects (Towill et al., 2007). While order-batching policies (e.g., Burbidge, 1987) may induce cyclical ordering patterns, retail seasonal smoothing policies may reflect a "rhythm" that is likely to repeat with each seasonal cycle (Parkany, 1961).

All of the above factors affect forecast accuracy of a new product extension as well. First, the bullwhip effect remains a significant influence on demand distortion between POS and order data. Utilizing order data as forecast basis incorporates potentially misleading information due to amplified variance. Second, suppliers must take into account of past retail order "rhythms" as well. It is highly unlikely for the retailer to order new product extensions in a similar pattern as other more well-established products within the same family. Therefore, the order "rhythms" might yield misleading customer demand forecast for new products.

However, certain product categories have clear seasonal peaks. For product lines competing under such categories, retail orders tend to reflect "rhythms" based on seasonal ordering policies, rather than idiosyncratic effects based on localized optimization in response to consumer demand. For example, to alleviate operational and capacity constraints, retailers tend to build inventory by placing steady orders with suppliers during low seasonal demand periods for rapid depletion when seasonal demand peaks. This retail ordering behavior is not likely to vary between seasonal products with similar consumer demand characteristics due to operational necessity. Thus, for seasonal products, order data for well-established products will likely be in a better position than its POS counterpart to assist suppliers in forecasting customer demand for new product extensions.

C. Data Science 3 – Incorporating the Marketing Mix in Forecasting

Extant marketing theory asserts that increased demand due to promotional activities related to a product tends to be short-lived and that increased demand will eventually revert back to a pre-shock level (Lautman & Pauwels 2009; Vakratsas & Ambler 1999). This notion had been further supported through similar studies. Wieringa & Horvath (2005) found that promotions in general provide only a short term increase in sales, which dissipate rapidly in the post-promotional period. Since most demand planning is performed for short term horizons, incorporating various factors of the marketing mix as variables exogenous to time series-based forecasting techniques can potentially increase forecast accuracy. In addition, this is likely to also allow marketers to evaluate the effectiveness of various marketing tools for certain products and product categories.

The logistics-marketing interface has a long stream of literature. The majority of this literature examines the cross-functional and cross-boundary impact of integrating logistics and marketing processes. Most frequently, measurements of antecedents to and effects of logistics-marketing integration are done in survey format. Although the literature recognizes many benefits, such as more effectively matching supply with demand, in incorporating marketing factors into logistics processes, very few studies validated these results using sales data.

In demand forecasting, a significant body of knowledge is built on specific methodologies and forecast models. They typically explore and examine performance differences among various forecast specifications and parameters. More recently, Williams and Waller (2010; 2011) introduce shared downstream demand signals as a forecast basis for retail orders. They show that downstream demand signals can significantly outperform retail order history in forecast accuracy. Main reason for the improvement is due to lower stochastic variance and superior information content of the downstream demand signal. Therefore, determinant of forecast accuracy include factors include the statistical properties of the forecast input, the specified forecast model, as well as any relevant forecast parameters.

Improved demand forecasts may be expected with the inclusion of the marketing mix. While downstream demand signals contain information that may potentially reveal consumer demand patterns, these patterns cannot be attributed to any specific causes. This is problematic because consumer response to marketing factors may change over time. The total demand variance can be segregated into stochastic and seasonal components. While the seasonal variance may be considered deterministic, the stochastic variance can be the result of exogenous causes. For example, the bullwhip literature notes that external influences such as managerial gaming behavior amplify stochastic demand variance (Lee et al., 1997). Since consumer demand is typically composed of base and marketing components (Lautman & Pauwels, 2009), incorporating exogenous variables such as the marketing mix may allow demand planners to utilize their estimated effects on consumer response to forecast future demand.

D. DATA SCIENCE 4 – BIG DATA IN SUPPLY CHAIN

In the retail supply chain, firms have been collecting massive amount of transaction data for the last several decades. Only recently did supply chain companies begin to truly leverage data to make decisions with increased precision and at a faster pace. Aided by rapid improvement to information technology, companies have further begun to amass other forms of data, ranging from consumer demographics, to social media, to geo-cache locations, all of which are collectively coined “Big Data”. Unlike the traditional transaction-based data, which can be conveniently structured based on location, category, time, and even customer, most forms of Big Data are unstructured. Without models specified a priori, some industry executives declared the death of traditional forms of strategic business leadership and advocated for correlation-driven decisions.

In an Aspen Institute conference on business applications of Big Data, participants noted that while many businesses are formed to cater to correlation-driven opportunities, many statistical oddities have resulted. The inherent danger in such an approach to business is that spurious relationships can be identified to mislead companies into devoting large amount of resources to business opportunities that really aren’t there. Furthermore, the expenses of processing such volumes of data tend to be quite high. Companies such as Amazon.com choose to automate this correlation-driven process, which resulted in unintentionally comical consequences, such as the “my TiVO thinks I’m gay” phenomenon (Bollier, 2010, p. 23).

Thus, in order for supply chain companies to leverage Big Data for greater decision-making, several questions must be answered. First, given that some pioneering companies such as Wal-Mart and Amazon had long utilized data to drive business decisions, how is the current movement of Big Data different from what managers traditionally had known? Second, despite statisticians’ warnings on spurious correlations, limited business successes can be readily observed. Thus, research should be undertaken to differentiate when correlation-driven decisions are appropriate as opposed to deductive reasoning, because clearly both methods of analysis can have positive impact on firm performance. And lastly, despite the benefits of data-driven decisions, it remains unclear as to what will ultimately be the main drivers of Big Data adoption. Identifying the internal and external drivers of Big Data's adoption and proper use can define a company's success.

E. SUPPLY CHAIN FINANCE 1 – EMPIRICAL EFFECT ON PERFORMANCE AND VALUATION

Supply chain performance had been positively linked with many business outcomes. They include measures ranging from operations to profitability. While empirical linkage between operations and profitability performance is well documented among various streams of literature, the supply chain literature relies primarily on surveys and interviews to document these linkages. However, few studies exist to verify that these linkages materialize into positive market response.

Publicly-traded companies may see their market values appreciate for various reasons. Most commonly, past performance measures such as sales growth, profitability, market share, and future performance factors such as product pipeline, pending patents, and expected market expansions. Considering the various documented positive impact of supply chain competence, market valuation should also reflect, to a degree, investors' understanding of a firm's supply chain capabilities. Corroborating with this intuition, Hendricks and Singhal (2005) found that supply chain disruptions can result in substantial harm to a firm's stock price.

Supply chain management is viewed by many scholars as a vital firm resource. For example, supply chain management is found to be a source of competitive advantage due to the value it adds to growth in the firm's top-line sales as well as bottom-line profit. Furthermore, it may also be considered sustainable and not substitutable because its success hinges on a tremendous amount of investment as well as top management commitment (Jin et al., 2013). Even after firms purchase the necessary physical assets for managing its supply chain, their effective and efficient use is often described as a capability (Allred et al., 2012), which is intangible and develops over time (Teece et al., 1997). Therefore, a firm's supply chain capabilities are not only dependent upon possessing the necessary assets, but also the ability to deploy them and to maximize their utility. Indeed, streams of literature exploring supply chain collaboration and integration treat these concepts as largely dynamic capabilities rather than tangible assets. But aside from Hendricks and Singhal (2005) and Ellinger et al. (2011), few studies explore the supply chain-finance interface.

Thus, while the SCM literature frequently identifies various benefits to firm performance, whether the market rewards firms for supply chain excellence is far less clear. For example, although Amazon.com is almost universally championed for its logistics and supply chain innovations today, during its early days the market frequently hammered it for incurring too much R&D expenses while depressing profitability. While limited research exists in examining supply chain ranking's impact on firm default risk (Ellinger et al., 2011) and supply chain disruption's effect on stock prices (Hendricks & Singhal, 2005), little evidence exists that the market actively recognizes and rewards firms for supply chain excellence.

A study can be developed through a combination of several sources of data. First, Gartner ranks the world's top supply chains annually. Although only the top 25 are ranked each year, rankings are available from 2004 to 2012 (except for 2007), which provides a small but decently-sized panel of observations to accommodate some control variables and effects. Many of these companies are publicly traded.

In terms of variables of interest, several supply chain-related variables have been developed over time. Rumyantsev and Netessine (2007) used several proxy variables based on financial data to examine classic inventory systems. Eroglu and Hofer (2011) developed an empirical leanness indicator to reflect lean inventory management strategies. Further, companies also use variables constructed using various financial data to represent factors such as cash-to-cash conversion cycle, gross margins return on inventory and others to capture supply chain management's impact on firm operations.

To measure market reaction, stock price is not entirely appropriate. A primary reason is because stock price is not a perfect measure of the overall investor sentiment on a firm's future performance. While a firm's stock price may fluctuate due to past and future performance factors, changes to the firm's assets may also influence stock price. Tobin (1969) argued that a firm's market value should be about equal to their replacement if only tangible assets were concerned. Any premium the market places over the total firm physical asset reflects investors' view on the value of the firm's intangible assets and capabilities. Since supply chain management is not usually explicitly measured as a form of firm asset nor can it be quantified as a firm capability, its impact can therefore be reflected in the firm's market value to total assets ratio, or Tobin's Q (Tobin, 1969).

Therefore, a research study could be formed to measure the various impacts of supply chain outcomes, as measured by variables representing operational performance, on the market value premium placed on the firm's physical assets, as measured by Tobin's Q, part of which may be theorized as a reflection of supply chain management premium.

Financial Crisis Extension

A primary function of supply chain capability is to grant firms the capability in mitigating negative market influences. Various theories, such as organizational modularity and ambidexterity also support the notion that firm flexibility and agility, which may be enhanced by supply chain management, allowing firms to quickly and effectively react to market shocks. The financial crises presented unprecedented degree of market uncertainty. Most companies were caught off guard by the sudden drop in demand. Therefore, the performance of supply chain leaders in the financial crisis is interesting to be examined as well using similar data.

F. SUPPLY CHAIN FINANCE 2 – EMPIRICAL EFFECT ON EXECUTIVE COMPENSATION

In the finance and accounting literature, a large stream of research exists on executive compensation. Often grounded in agency theory, firm performance should determine executive compensation because stock options and bonus incentives for senior management should align their interest with the shareholders' interest—outcome-based contracts.

In practice, executive compensation may be divided in to several categories. They typically include salaries, short-term bonuses based on performance, long-term incentive systems such as stock options, fringe benefits. In addition, many executives negotiate "golden parachutes," which are pre-negotiated severance pay, often in exceedingly high amounts, in the event the executive is forced to resign from the company.

Salaries are the immediate and set amount of compensation paid to executives for their services rendered to the company. Short-term bonuses are often based on immediate company performance goals and are driven by formula reflective of financial performance for the previous fiscal period. Long-term incentive systems are more complex and by far the most studied compensation scheme in finance, accounting, and management fields. Jensen and Murphy (1990) theorized and empirically supported that restricted stock options (i.e., cannot be exercised until 3 to 5 years after issuance) aligns managers' interest in the company's long term performance in maximizing shareholder wealth, thereby mitigating agency costs. On the other hand, empirical evidence also suggest that long term stock options contributed to managerial incentive to accounting manipulation scandals as well as initiating stock buyback programs just before exercising vested options (Bebchuk & Fried, 2006). Yet little literature exists on the relationship between operational performance and executive compensation.

In the supply chain literature, executive compensation studies are almost non-existent, though some studies tangentially related to this topic have explored incentive systems and non-financial performance measures. Banker, Potter, and Srinivasan (2000) examined how the inclusion of nonfinancial performance measures in an incentive plan impact firm financial performance. They find that nonfinancial measures such as customer satisfaction and capacity utilization are significant contributors to financial performance. Furthermore, the inclusion of these nonfinancial measures as part of managerial incentive system also improved financial performance.

Based on survey research, an abundance of case studies as well as survey research indicates that supply chain excellence positively contributes to firm performance. Moreover, executive commitment to dedicating adequate amount of resources toward developing robust supply chain directives is a vital component of efficient and effective supply chain management (Jin et al., 2013). While abundant survey research and case studies exist to support supply chain management's benefit to firm performance, executive commitment is first and foremost motivated by the proper incentive structure. Considering the role of incentives as a tool to drive executive decisions, insight into whether superior supply chain performance contributes to executive compensation can provide substantial justification to command greater executive commitment to dedicate resources to improve supply chain management.

Yet, little knowledge exists in supply chain management with regard to how the purported firm performance benefits translate to executive rewards. This is partly due to an overall lack of acknowledgement in industry that supply chain management warrants its own executive oversight. Often, supply chain managers are grouped under operations or marketing. While a current trend among board structure is to assign supply chain, transportation, and logistics functions to a dedicated executive, much is unknown if current executives are already seeing financial benefits from supply chain performance.

In extant literature, supply chain management is positively linked with firm financial performance by increasing sales, lowering costs, improving customer satisfaction and operational performance, and employee satisfaction. All of these factors are also linked to overall financial performance benefits. Since executive compensation is based on a combination of both short term and long term firm performance factors, a logical extension can be made that supply chain management may also increase executive compensation by improving the above factors. Thus, a highly relevant question central to bringing executive attention to recognize the importance of supply chain excellence is to examine how nonfinancial performance factors that are often linked to supply chain excellence affect executive compensation.

Specifically, the question could be answered in several stages. First, what aspect of firm operational performance is directly influenced by supply chain excellence? Inventory management is certainly a highly relevant factor. Perhaps asset utilization in certain industries may also be particularly influenced by supply chain excellence. A thorough literature review may yield more insight. Second, a list of top supply chain companies may be identified from Gartner's Top 25 Supply Chain Companies report. Third, public financial data from CompuStat may be used to construct measurements of operational performance based on the factors identified based on literature review. Lastly, executive compensation data can be obtained through Execucomp. Empirical models could then be developed based on the nonfinancial factors.

Data analysis can be conducted in several ways. First, since Gartner's list spans many different industries, therefore comparing between companies might not be valid. Second, the list was initiated in 2004. The most recently available list is for 2012. Furthermore, not all listed companies are publicly traded. Therefore, limited degrees of freedom also substantially limit the statistical power of the model as well as the ability to accommodate statistical controls. Perhaps a prudent analysis would be to identify the six digit NAICS code for each of the publicly-traded company for each year, and obtain financial statistics for all of the company's competitors to calculate necessary operational performance proxy variables. Finally, each company's operational performance variables can be normalized to its respective industry in order to make measures comparable across all companies in the sample.

G. STRATEGIC RETAIL 1 – EDLP VERSUS HI-LO PRICE MANAGEMENT

Retailers generally use one of two pricing models: everyday low price (EDLP) and high-low (Hi-Lo). For those retailers competing under EDLP, they tend to avoid heavy price reductions with which to draw customers. Instead, most of their products are priced at levels below most of their competitors in similar retail formats. EDLP advocates claim that such a pricing approach alleviates price anxieties among customers and provides reassurance that on average, they will spend less money than shopping at a competing retailer. On the other hand, Hi-Lo retailers are generally not concerned with having predominantly low prices. Instead, they rely on short-term temporary price reductions, typically on a weekly basis, to draw consumers in. The discounted items are usually sold for little gross profit, if not as loss leaders altogether. In turn, these retailers hope that the consumers would purchase higher margined items for either convenience or impulse to make up for the price promotions.

Recently, JCPenney became the latest high profile retailer to experience a failed transition from a Hi-Lo model to EDLP. Though many reasons have been speculated as contributors to this failed transition, there is one common criticism from among all analysts and industry experts. JCPenney's customers simply did not respond to the value proposition offered by the EDLP model. While its ex-CEO who initiated this transition process clearly stated that the EDLP model offers no gimmicks, just honest pricing, consumers appear to be less enthused. In addition, any potential cost benefit due to the transition away from significant price variations was masked by unsold inventories of merchandise.

From a retail operations perspective, EDLP offers substantial benefits. Without price variation, demand variance becomes significantly lower (Lee et al., 1997). With lower demand variance, retail operations-planning also becomes easier as forecasting is likely to be more accurate. The smoothed demand brings benefits associated with lowered contingent resources such as transportation, storage, and labor capacities. Furthermore, due to lowered variance of demand, future forecast is likely to be more accurate as well.

However, consumer behavior theories argue that price promotions offers incentive for consumers to make the immediate purchase. Literature had also shown that more frequent purchases translate readily into more frequent consumption, thereby increasing both immediate and future consumption altogether. In addition, Hi-Lo retailers can also create significant enthusiasm among shoppers with attention-grabbing prices. The downside to Hi-Lo retailing, of course, is the fact that demand due to price promotion is very difficult to forecast. Inadequate inventory levels lead to out-of-stock, which in turn causes delayed purchase, consumer ill-will, and potential loss of not only the immediate sale but also an entire basket. In worst cases, the jilted consumer may decide to not return to the retailer altogether.

Given the many advantages and disadvantage associated with both pricing models, the literature remains unclear as to which format is more appropriate. On the one hand, retailers such as Wal-Mart and Costco built significant competitive advantage on the market by offering EDLP. On the other hand, highly successful stories among Hi-Lo retailers further reinforced the perception that price promotion is a valuable tool for retail sales. Furthermore, the transition from Hi-Lo to EDLP appears to be fraught with difficulties and anticipated consequences. Therefore, greater understanding of drivers of EDLP and Hi-Lo pricing models as well as their advantages can have substantial impact on retail operations management.

H. STRATEGIC RETAIL 2 – PRODUCT VARIETY AND RETAIL PERFORMANCE

In Wal-Mart's relentless drive to lower supply chain costs, consequences of its recent SKU reduction (Roberts & Berg, 2012) suggest that its consumers have grown so accustomed to a wide variety of choices at low prices that they were unwilling to accept tradeoffs—lower prices but fewer choices. The press literature documents many stories of how Wal-Mart is being squeezed by retailers that emphasize smaller selection but cheaper prices. These retailers run the gamut from big warehouse clubs such as Costco to small stores such as Dollar General. On the other hand, stores such as HEB carry a tremendous variety with a heavy emphasis on grocery—traditionally the retail segment with the lowest margins—and manage to remain profitable at competitive prices. Thus it appears that Wal-Mart, having been dominant for so long, is being squeezed from both ends of the competitive spectrum.

Two competing theories in product variety exist from two fields. In logistics, an optimization approach argues that an optimal product variety provides the greatest firm performance. With a single product, demand can be highly variable and random. As product variety increases, demand peaks and troughs may be smoothed depending on the degree of correlation between products. As a result, demand planning for transportation, storage, and labor capacities become easier due to reduced overall variance relative to the mean. However, increase in product variety also results in greater coordination costs. Even as demand planning for capacity becomes easier with increased product variety, increased coordination costs can ultimately erode product variety's benefits. As a result, the logistics perspective argues that the relationship between product variety and performance is in an inverted-U shape.

On the other hand, the marketing literature argues that product variety should be either very low or very high. That is, the relationship between product variety and firm performance is U-shaped. At the lowest end, low product variety allows simpler operations and replenishment as well as bulk purchase discounts. With bulk discounts as well as overall lower operations costs, retailers can position themselves as cost leaders. Considering that much of the cost benefits are due to this particular operating format, its price advantage is likely persistent against competitors that do not attempt to replicate this strategy. Notable examples of success under this format include discount grocers such as Aldi, as well as most discount warehouse clubs such as Costco. At the highest end, high product variety provides greater service to draw customers into the store. By providing choice, retailers are able to provide superior choices to consumers by holding multiple SKUs that serve similar functionality, often with only minor differences such as color, taste, and scent. Due to the complexity involved in making such arrangements, costs are often high for these retailers. Therefore they cannot always compete on price against retailers that provide limited selections. On the other hand, these retailers are also relatively insulated from competitors that cannot compete on selection.

In retail management, many retailers often attempt to re-tool their product selection to balance the need for providing consumers with the choices that they want, while holding costs low. Only rare exceptions such as Wal-Mart compete on both variety and price. Yet even Wal-Mart's failed SKU rationalization program suggests that the relationship among variety, price, and performance is often delicate and not very well defined. Furthermore, two competing theories on the functional form between SKU variety and retail performance indicates that further understanding is required. Therefore, determining the strategic benefit of SKU variety has important implications for retail management. Is SKU-optimization ultimately a myopic or strategically competitive goal?

I. MISCELLENEOUS 1 – AGENCY IN SUPPLY CHAIN

In the classic agency theory, two actors establish a relationship in which one is the agent and the other is the principal. The agent is charged with acting in the best interest of the principal, who in turn rewards the agent for the services rendered. Supply chain partners are unique in that both actors carry dual-roles of the principle and the agent.

In a typical retail supply chain partnership, the retailer may be viewed as the agent for the supplier in a sense that the retailer sells the products for its suppliers. With greater retail market penetration, the supplier can enjoy greater sales. Thus, a revenue and profit incentive had traditionally aligned the goals of both suppliers and the retailers. However, as the supplier becomes increasingly dependent on the retailer for purposes other than a simple sales-based relationship, agency conflicts begin to arise.

Recent advances in information technology had enabled suppliers to adopt increasingly more sophisticated demand management tactics. Whereas supply chain management prior to the internet era generally emphasized efficient downstream physical distribution, the current dominant strategies include information flow back upstream as well. A primary benefit to bi-directional flow is that suppliers can generate more accurate demand forecast to further increase the efficiency and effectiveness of downstream physical distribution. In addition to improved service levels, retailers may also expect the suppliers to pass certain supply chain savings down. In that sense, the previous principal-agent relationship had also become effectively bi-directional. The supplier depends on the retailer for executing store-level sales while the retailer depends on suppliers to properly utilize sales information and at the same time not using it to help other competing retailers.

J. MISCELLANEOUS 2 – EMPATHETIC CONCERN AND MACHIAVELLIANISM IN SUPPLY CHAIN ORIENTATION

Modern supply chain management strategies often calls for supply chain partners to make decisions based on shared demand and supply information for synchronized activities. These strategies are seldom unstructured and informal. Instead, industry initiatives such as CPFR, S&OP, VMI, and other programs (e.g., Waller et al., 1999; Lapide, 2007) provide blueprints for such endeavors. As part of these strategies, sensitive information are frequently passed between supply chain partners through automated processes, which are enabled by interorganizational systems that provide a common platform (Bendoly & Cotteleer, 2008). Under this framework, the buyer, who typically occupies a position downstream along the supply chain, would transmit demand signals upstream to its supplier. The supplier, in turn, would leverage the additional information for demand planning purposes (Williams & Waller, 2010; 2011). Many benefits associated with such integrated supply chain management programs have been anecdotally described by participants and also empirically validated by scholars (e.g., Lee et al., 1997; Wu and Katok, 2006; Allred et al., 2011).

Despite the widely publicized benefits, stories of failures remain abundant as well. Collaborative management efforts had been repeatedly criticized as being more rhetoric than reality (Fawcett & Magnan, 2002), that its results are frequently disappointing (Sabbath and Fontanella, 2002). While many firms attempted to reap the benefits proposed by information sharing as a form of integration, results are not universally positive (Daugherty et al., 2006). Furthermore, the relationship between information sharing and performance is not universally positive (Fabbes-Costes & Jahre, 2007) and may instead be linearly positive and has an optimum level, beyond which performance begins to decrease (Fabbes-Costes & Jahre, 2008). More recently, Jin et al. (2013) found that while many firms continued to advance information sharing and integration with supply chain partners, others regressed with some dropping such efforts altogether.

Many attempts have been made by scholars to identify specific determinants of information sharing's success. Specifically, information sharing improves managerial decision-making through greater visibility (Barratt & Oke, 2007). As a result, information sharing improves both internal and external coordination efforts (Mentzer et al., 2004), customer service quality (Lee and Whang, 2000), increase forecast accuracy (Williams & Waller, 2010; 2011), and lower agency conflicts (Nyaga et al., 2007) as well as overall supply chain costs (Datta et al., 2007). In order for the above benefits to materialize, the shared information should be timely and relevant (Kaipia & Hartiala, 2006). However, most of the shared information is considered trade secrets that can be used against firms that shared it. As a result, many firms remain reticent to share data exactly as needed by the supplier to make effective decisions.

Supply chain orientation is frequently cited as an antecedent for successful integration efforts to coordinate and synchronize activities (Min & Mentzer, 2004). In order for firms to acquire a supply chain orientation, psychometric qualities such as trust and commitment are vital. Although interfirm relations in the supply chain literature are typically measured at the organizational level, specific transactions occur and relationships are built between individual company representatives. Moreover, reliance on individual relationships affects both interfunctional as well as interfirm integration efforts (Fawcett & Magnan, 2002; Barratt & Oke, 2007; Cousins et al., 2006). Therefore, the road that firms expect to take directly from integration to supply chain performance benefits is more frequently serpentine rather than direct as agency conflicts abound (Rungtusanatham et al., 2007).

Agency conflicts exist partially due to misaligned incentive systems as well as information asymmetry between two parties both seeking self-interest (Jensen & Meckling, 1976). For example, Lee et al. (1997) illustrated a prominent example of misaligned incentive systems in the supply chain, in which a supplier wants to ensure broad geographic reach while a buyer wants to maximize sales. As a result, the supplier attempts to ration quantities of shipment while the buyer orders at quantities higher than forecasted in order to secure more products. A prominent consequence of these actions, distorted demand signals, can be observed in a wide range of industries empirically (e.g., Forrester, 1961; Blanchard, 1983; Eichenbaum, 1989; Fransoo & Wouters, 2000; Waller et al., 2008; Bray & Mendelson, 2012) and also replicated in a number of behavioral experiments (e.g., Sterman, 1989; Sterman, 1992; Cantor & Katok, 2012; Tokar et al., 2012).

While behavioral conflicts may be ameliorated through contract designs, effective supply chain relationships should be based on relational integration rather than explicitly governing contingent outcomes. Hence, having confidence in the partner's propensity in making the optimal decision based on mutual rather than self-interest is a vital building block to mitigate conflicts due to information asymmetry and misaligned incentive systems.

Reliance on the partnering manager to "do the right thing" on behalf of both firms makes an inherent humanist assumption. By making such an assumption, managers may believe that their counterpart at a partnering supply chain firm would act based on the core values of their joint partnership and not engage in actions based solely on self-interest. Indeed, many scale items in supply chain literature place specific emphasis on the degree of trust placed by survey respondents on partnering firm to make decisions that are mutually beneficial (e.g., Zaheer et al., 1998; Cai et al., 2010). Under this assumption, results from surveys largely support the contention that trust and commitment both lead to supply chain orientation, which in turn enables superior supply chain performance as measured by various outcomes.

However, anecdotal stories of the opposite are pervasive in case studies as well as the press literature. For example, many suppliers to major retailers elucidate a coercive dependence relationship (e.g., Bloom & Perry, 2001). Classic cases such as Vlasic Pickles failure (Fishman, 2006) demonstrate how a humanist assumption in supply chain management is not necessarily appropriate. More recently, a highly publicized lawsuit over alleged breach of contract involving executive leaderships from Macy's, JCPenney, and Martha Stewart (D'Innocenzio, 2013), who was counted by Terry Lundgren, the CEO of Macy's, as a personal friend (Tuttle, 2013), further highlight the tenuous relationship between representatives from supply chain partners.

Clearly, a humanist assumption in supply chain management does not necessarily provide the entire picture of antecedents to supply chain collaboration. Therefore, while information sharing may be considered to be positively related to supply chain performance, personal characteristics of individual company representatives contributes to the ultimate effectiveness of such linkages.

Following the humanist perspective, a principle component of trust and commitment is empathetic concern. Empathetic concern describes the capability for an individual to have a positive regard or a non-fleeting concern for the other party (Chismar, 1988). Greater empathetic concern allows individual managers to take the position of their business partner, to recognize the value proposition as well as needs and incorporate their partner's concerns into their own decision-making process. In supply chain management, collaborative tactics such as information sharing is often described as being beneficial primarily to firms residing upstream along the supply chain (Dukes et al., 2009). Under such unevenly distributed value appropriation, the humanist manager from the firm downstream would empathetically recognize that there is substantial gain to be had by his counterpart from the firm upstream, and vice versa. When both partners have the same realization, the humanist perspective suggests that despite the uneven distribution of benefits, both partners would seek to make the ideal decision to realize maximum value for the supply chain.

On the other hand, recent studies had found that many managers—especially those residing the upper echelons of organizations—tend to display psychopathic traits (e.g., Newby, 2005) that prevent empathetic concerns (Howard & McCullagh, 2007). Under such conditions, value maximization shifts from the system-level to individual-level. As a result, classic gaming behaviors would occur, as described by the negative consequences on supply chain performance due to self-interest (Lee et al., 1997; Rungtusanatham et al., 2007).

K. REFERENCES

Allred, C.R., Fawcett, S.E., Wallin, C., & Magnan, G.M. (2012). A dynamic collaboration capability as a source of competitive advantage. *Decision Sciences*, 42, 129-161.

Armstrong, G., & Kotler, P. (2006). *Marketing: An Introduction*, Upper Saddle River, New Jersey: Pearson Prentice Hall.

Banker, R.D., Potter, G., & Srinivasan, D. (2000). An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review*, 75 (1), 65-92.

Barratt, M., & Oke, A. (2007). Antecedents of supply chain visibility in retail supply chains: a resource-based theory perspective. *Journal of Operations Management*, 25, 1217-1233.

Bebchuk, L.A., & Fried, J.M. (2006). Pay without performance: overview of the issues. *Academy of Management Perspectives*, 20 (1), 5-24.

Bendoly, E., & Cotteleer, M.J. (2008). Understanding behavioral sources of process variation following enterprise system deployment. *Journal of Operations Management*, 26 (1), 23-44.

Blanchard, O. J. (1983). The Production and inventory behavior of the American automobile industry. *Journal of Political Economy*, 91 (3), 365-400.

Bloom, P.N., & Perry, V.G. (2001). Retailer power and supplier welfare: the case of Wal-Mart. *Journal of Retailing*, 77 (3), 379-396.

Bollier, D. (2010*) The Promise and Peril of Big Data*. The Aspen Institute, Washington DC.

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

Burbidge, J.L. (1981). The new approach to production. *Production Engineer*, 40 (12), 769-784.

Cai, S., Jun, M., & Yang, Z. (2010). Implementing supply chain information integration in China: The role of institutional forces and trust. *Journal of Operations Management*, 28 (3), 257-268.

Cantor, D.E., & Katok, E. (2012). Production smoothing in a serial supply chain: A laboratory investigation. *Transportation Research Part E*, 48, 781-794.

Chismar, D., (1988). Empathy and sympathy: the important difference. *Journal of Value Inquiry*, 22, 257-266.

Conrad, S.A. (1976). Sales data and the estimation of demand. *Operational Research Quarterly*, 27 (1), 123-127.

Cousins, P.D., Handfield, R.B., Lawson, B., & Petersen, K.J. (2006). Creating supply chain relational capital: The impact of formal and informal socialization processes. Journal of *Operations Management*, 24 (6), 851-863.

Datta, S., Granger, C.W.J., & Gibbs, T. (2007). Management of supply chain: an alternative modeling technique for forecasting. *Journal of the Operational Research Society*, 58, 1459-1469.

Daugherty, P.J., Richey, R.G., Roath, A.S., Min, S., Chen, H., Arndt, A.D., & Genchev, S.E. (2006). Is collaboration paying off for firms? *Business Horizons*, 49, 61-70.

Desai, P.S. (2001). Quality segmentation in spatial markets: when does cannibalization affect product line design? *Marketing Science*, 20 (3), 265-283.

D'Innocenzio, A. (2013). Martha Strewart betrayed alliance, Macy CEO testifies. *The Boston Globe*, Retrieved from http://www.bostonglobe.com/business/2013/02/26/macy-ceo-testifies-fight-over-martha-stewart/e0z0tgi2yfYjXrOWf2sWXM/story.html

Dukes, A.J., Geylani, T., & Srinivasan, K. (2009). Strategic assortment reduction by a dominant retailer. *Marketing Science*, 28 (2), 309-319.

Eichenbaum, M.S. (1989). Some empirical evidence of the production level and production cost smoothing models of inventory investment. *American Economics Review*, 79 (4), 853-864.

Ellinger, A.E., Natarajarathinam, M., Adams, F.G., Gray, B., Hofman, D., & O'Marah, K. (2011). Supply chain management competency and firm financial success. *Journal of Business Logistics*, 32 (3), 214-226.

Eroglu, C., & Hofer, C. (2011). Lean, leaner, too lean? The inventory-performance link revisited. *Journal of Operations Management*, 29 (4), 356-369.

Fabbe-Costes, N., & Jahre, M. (2007). Supply chain integration gives better performance—the emperor's new suit? *International Journal of Physical Distribution and Logistics Management*, 37, 835-855.

Fabbe-Costes, N., & Jahre, M. (2008). Supply chain integration and performance: a review of the evidence. *International Journal of Logistics Management*, 19, 130-154.

Fawcett, S.E., & Magnan, G.M. (2002). The rhetoric and reality of supply chain integration. *International Journal of Physical Distribution and Logistics Management*, 32 (5), 339-361.

Fishman, C. (2006). The Wal-Mart effect and a decent society: Who knew shopping was so important? *Academy of Management Perspectives*, 20 (3), 6-25.

Forrester, J.W. (1961). *Industrial Dynamics*. Productivity Press, Portland OR.

Frank, R.E., Massy, W.F., & Wind, Y. (1972). *Marketing Segmentation*. Englewood Cliffs, New Jersey: Prentice Hall.

Fransoo, J.C., & Wouters, M.J.F. (2000). Measuring the bullwhip effect in the supply chain. *Supply Chain Management: An International Journal*, 5(2), 78-89.

Hendricks, K.B., & Singhal, V.R. (2005). An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm. *Production and Operations Management*, 14 (1), 35-52.

Howard, R., & McCullagh, P. (2007). Neuroaffactive processing in criminal psychopaths: Brain event-related potentials reveal task-specific anomalies. *Journal of Personality Disorders*, 21 (3), 322-339.

Huh, W.T., Janakiraman, G., Muckstadt, J.A., & Rusmevichientong, P. (2009). An adaptive algorithm for finding the optimal base-stock policy in lost sales inventory systems with censored demand. *Mathematics of Operations Research*, 34 (2), 397-416.

Jensen, M.C., & Meckling, W.H. (1976). Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3 (4), 305-360.

Jensen, M.C., & Murphy, K.J. (1990). Performance pay and top management incentives. *Journal of Political Economy*, 98, 225-264.

Jin, Y., Fawcett, A.M., & Fawcett, S.E. (2013). Awareness is not enough: commitment and performance implications of supply chain integration. *International Journal of Physical Distribution & Logistics Management*, 43 (3), 205-230.

Kaipia, R., & Hartiala, H. (2006), Information-sharing in supply chains: five proposals on how to proceed. *International Journal of Logistics Management*, 17 (3), 377-393.

Lapide, L. (2007). Sales and operations planning mindsets. *Journal of Business Forecasting*, 26, 21-31.

Lautman, M.R., & Pauwels, K. (2009). What is important? Identifying metrics that matters. *Journal of Advertising Research*, 49 (3), 339-359.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43 (4), 546-559.

Lee, H.L., & Whang, S. (2000). Information sharing in a supply chain. *International Journal of Manufacturing Technology and Management*, 1 (1), 79-93.

Mentzer, J.T., Min, S., & Bobbitt, L.M. (2004). Toward a unified theory of logistics.

Metters, R. (1997), Quantifying the bullwhip effect in supply chains. Journal of Operations Management, 15 (2), 89-100.

Meyer, M.H., Tertzakian, P., & Utterback, J.M. (1997). Metrics for managing research and development in the context of the product family. *Management Science*, 43 (1), 88-105.

Min, S., & Mentzer, J.T. (2004). Developing and measuring supply chain management concepts. *Journal of Business Logistics*, 25 (1), 63-99.

Moorthy, K.S. (1984). Marketing segmentation, self-selection, and product line design. *Marketing Science*, 2 (Fall), 288-310.

Nahmias, S., & Smith, A. (1994). Optimizing inventory levels in a two echelon retailer system with partial lost sales. *Management Science,* 40 (5), 582-596.

Newby, J. (2005). Corporate psychopaths, in *Catalyst*, Australia.

Nyaga, G.N., Closs, D.J., Rodrigues, A., & Calantone, R.J. (2007). The impact of demand uncertainty and configuration capacity on customer service performance in a configure-to-order environment. *Journal of Business Logistics*, 28 (2), 83-104.

Parkany, J. (1961), A new approach to sales forecasting and production scheduling. *Journal of Marketing*, 25 (3), 14-21.

Roberts, B., & Berg, N. (2012), *Walmart: Key Insights and Practical Lessons from the World's Largest Retailer,* Philadelphia, Pennsylvania: Kogan Page Limited.

Rumyantsev, S., & Netessine, S. (2007). What can be learned from classical inventory models? A cross-industry exploratory investigation. *Manufacturing and Service Operations Management*, 9 (4), 409-429.

Rungtusanatham, M., Rabinovich, E., Ashenbaum, B., & Wallin, C. (2007). Vendor-owned inventory management arrangements in retail: An agency theory perspective. *Journal of Business Logistics*, 28 (1), 111-135.

Sabbath, R., & Fontanella, J. (2002). The unfulfilled promise of supply chain collaboration. *Supply Chain Management Review*, 6 (4), 24-29.

Smith, W.R. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of Marketing*, 21 (1), 3-8.

Solomon, M.R., Marshall, G.W., & Stuart, E.W. (2008). *Real People, Real Choices, Real Stories*. Upper Saddle River, New Jersey: Pearson Publishing.

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

Teece, D.J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18 (7), 509-533.

Tobin, J. (1969). A general equilibrium approach to monetary theory. *Journal of Money, Credit, and Banking*, 1 (1), 15-29.

Tokar, T., Aloysius, J.A., & Waller, M.A. (2012). Supply chain inventory replenishment: the debiasing effect of declarative knowledge. *Decision Sciences*, 43 (3), 525-546.

Towill, D.R., Zhou, L., & Disney, S.M. (2007). Reducing the bullwhip effect: looking through the appropriate lens. *International Journal of Production Economics*, 108, 444-453.

Tuttle, B. (2013). Macy’s CEO: ‘Shocked… sick to my stomach’ after betrayal by former ‘friend’ Martha Stewart. *Time Magazine*, Retrieved from http://business.time.com/2013/02/26/macys-ceo-shocked-sick-to-my-stomach-after-betrayal-by-former-friend-martha-stewart/

Vakratsas, D., & Ambler, T. (1999). How advertising works: what do we really know? *Journal of Marketing*, 63 (1), 25-43.

Waller, M.A., Johnson, E.M., & Davis, T. (1999). Vendor-managed inventory in the retail supply chain. Journal of Business Logistics, 20 (1), 183-203.

Waller, M.A., Williams, B.D., & Eroglu, C. (2008). Hidden effects of variable order review intervals in inventory control. *International Journal of Physical Distribution and Logistics Management*, 38 *3), 244-258.

Wieringa, J.E., & Horvath, C. (2005). Computing level-impulse response of log-specified VAR systems. *International Journal of Forecasting*, 21 (2), 279-289.

Williams, B.D., & Waller, M.A. (2010), Creating order forecasts: point-of-sale or order history? *Journal of Business Logistics*, 31, 231-251.

Williams, B.D., & Waller, M.A. (2011), Top-down vs. bottom-up demand forecasts: the value of shared point-of-sale data in the retail supply chain. *Journal of Business Logistics*, 32, 17-26.

Wu, Y., & Katok, E. (2006). Learning, communication, and the bullwhip effect. *Journal of Operations Management*, 24 (6), 839-850.

Zaheer, A., McEvily, B., & Perrone, V. (1998). Does trust matter? Exploring the effects of interorganizational and interpersonal trust on performance. *Organization Science*, 9 (2), 141-159.