AN EXTENSIVE EMPIRICAL ANALYSIS OF THE BULLWHIP EFFECT IN SUPPLY CHAINS

(TEDARİK ZİNCİRLERİNDE KAMÇI ETKİSİNİN KAPSAMLI EMPİRİK ANALİZİ)

by

Alper Yasin SARICIOĞLU, M.S.

Thesis

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

in

INDUSTRIAL ENGINEERING

in the

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

of

GALATASARAY UNIVERSITY

Supervisor: Prof. Dr. Müjde GENEVOIS

January 2025

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my thesis supervisor, Prof. Dr. Müjde GENEVOIS, whose unwavering encouragement and invaluable guidance have been instrumental throughout this journey. I also extend my sincere thanks to the monitoring committee members, Assoc. Prof. Dr. Ethem ÇANAKOĞLU and Assist. Prof. Dr. İlke BEREKETLİ ZAFEIRAKOPOULOS, as well as my friend Dr. Michele CEDOLIN, for their valuable contributions and insightful feedback. Additionally, I appreciate the time and feedback provided by the examining committee members, Prof. Dr. Başar ÖZTAYŞİ and Prof. Dr. A. Çağrı TOLGA, whose evaluations have helped refine this work.

I am profoundly grateful to my wife, Deniz, and my little son İlber Bora for their patience and understanding. Their constant support and steadfast presence provided me with the foundation needed to pursue this work with determination, and I could not have completed this dissertation without them.

Finally, I owe a special debt of gratitude to my parents, Hayati and Nedime SARICIOĞLU, whose unconditional love and encouragement have been crucial in every step of my academic journey. Their support has been a pillar of strength, and this achievement would not have been possible without them.

January 2025

Alper SARICIOĞLU

TABLE OF CONTENTS

LIST OF SYMBOLS

ADF	: Augmented Dickey-Fuller
ARMA	: Autoregressive Moving Average
BEA	: Bureau of Economic Analysis
BWE	: Bullwhip Effect
FMCG	: Fast-Moving Consumer Goods
LOESS	: Locally Estimated Scatterplot Smoothing
LSTM	: Long Short-Term Memory
M3	: Manufacturers' Shipments, Inventories, and Orders
MARTS	: Monthly Retail Trade Survey
MWTS	: Monthly Wholesale Trade Survey
NAICS	: North American Industry Classification System
OEE	: Overall Equipment Efficiency
OLS	: Ordinary Least Squares
RBWE	: Reverse Bullwhip Effect
RNN	: Recurrent Neural Network
SARIMA	: Seasonal Autoregressive Integrated Moving Average
SKU	: Stock Keeping Unit
USCMA	: United States-Mexico-Canada Agreement
VMI	: Vendor Managed Inventory
WITS	: World Integrated Trade Solution

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The Bullwhip Effect (BWE), defined as the amplification of demand variability as orders move upstream in a supply chain, remains a fundamental challenge in supply chain management. While extensive theoretical studies have explored its causes and consequences, empirical analyses at the industry level remain limited. Many assumptions regarding BWE's effects on inventory management, production efficiency, and labor stability have not been thoroughly tested across different industries. This dissertation addresses this gap by analyzing industry-level data to examine the key drivers of BWE, assess its impact during systemic disruptions such as COVID-19, and propose advanced forecasting techniques to mitigate its effects.

This study first evaluates the adverse effects of BWE across industries by analyzing key supply chain performance indicators, including inventory levels, backorders, overtime hours, job losses, and capacity utilization. The findings reveal that industries do not respond to BWE uniformly. While conventional theories suggest that inventory smoothing reduces BWE, empirical evidence indicates that this is not consistently observed across industries. Similarly, backordering, often assumed to amplify demand fluctuations, does not significantly contribute to BWE, whereas job loss rates exhibit a stronger correlation. These findings highlight that traditional assumptions about BWE's effects do not hold universally and that industry-specific approaches are required for effective mitigation.

The second part of the research examines the role of importation in exacerbating BWE. While extended lead times are often blamed for amplifying demand variability, this study finds that lead time alone does not necessarily drive BWE. Instead, factors such as information distortion, supply chain rigidity, and demand signal misinterpretation play more significant roles. This suggests that rather than prioritizing nearshoring or reshoring

as a blanket solution, firms should focus on strengthening supply chain collaboration and real-time information sharing to mitigate demand amplification.

The dissertation also investigates the impact of the COVID-19 pandemic on BWE using industry-level data. By comparing pre-pandemic and pandemic-era data, the findings demonstrate that BWE significantly intensified during the crisis, particularly in industries with high demand uncertainty and limited supply chain flexibility. These results underscore the importance of proactive risk management strategies to enhance supply chain resilience in future disruptions.

To address these challenges, the final part of this research explores the effectiveness of advanced forecasting methods in mitigating BWE. Traditional forecasting techniques often fail to capture nonlinear demand patterns, contributing to demand signal distortions. This study demonstrates that machine learning models, specifically Long Short-Term Memory (LSTM) networks and LightGBM, significantly improve forecast accuracy, reduce demand amplification, and enhance inventory stability. The findings suggest that integrating AI-driven forecasting methods into supply chain management practices can mitigate BWE and improve operational efficiency.

This dissertation makes several contributions to supply chain management literature by empirically testing the impact of BWE across industries, mapping its key drivers, and evaluating the effectiveness of machine learning-based forecasting methods. The results challenge traditional assumptions about BWE, demonstrating that industries respond differently to its effects and that lead time is not the sole determinant of demand amplification. The findings emphasize the importance of real-time data sharing, digitalization, and collaborative forecasting rather than solely focusing on structural supply chain adjustments. By providing practical recommendations for mitigating BWE, this research offers valuable insights for both academia and industry professionals, guiding the development of more resilient supply chains in an increasingly volatile global market.

ÖZET

Bu tez, Kamçı Etkisi'nin tedarik zinciri performansı üzerindeki etkilerini incelemektedir. Kamçı Etkisi, siparişlerin tedarik zinciri boyunca yukarı yönlü hareket ederken talep değişkenliğinin artması olarak tanımlanmaktadır. Literatürde bu etkinin nedenleri ve sonuçları geniş çapta tartışılmış olsa da endüstri düzeyinde yapılan ampirik çalışmalar sınırlıdır. Özellikle stok yönetimi, üretim verimliliği ve iş gücü dinamikleri üzerindeki etkileri her sektör için yeterince incelenmemiştir. Bu tez, endüstri verilerini analiz ederek Kamçı Etkisi'nin temel belirleyicilerini ortaya koymayı, kriz dönemlerinde nasıl değiştiğini değerlendirmeyi ve bu etkiyi azaltmaya yönelik gelişmiş tahmin yöntemleri sunmayı amaçlamaktadır.

Çalışmada ilk olarak, Kamçı Etkisi'nin farklı sektörlerde yarattığı olumsuz etkiler analiz edilmiştir. Stok seviyeleri, geri sipariş oranları, fazla mesai saatleri, iş gücü kayıpları ve kapasite kullanım oranları gibi tedarik zinciri performans göstergeleri incelenmiştir. Elde edilen bulgular, tüm sektörlerin Kamçı Etkisi'nden aynı şekilde etkilenmediğini göstermektedir. Stok yönetiminin her zaman talep dalgalanmalarını azaltmadığı, geri sipariş uygulamalarının ise beklenildiği gibi Kamçı Etkisi'ni artırmadığı görülmüştür. Ancak iş gücü kayıplarının Kamçı Etkisi ile güçlü bir ilişki içinde olduğu belirlenmiştir. Bu sonuçlar, Kamçı Etkisi'nin etkileri hakkında genel geçer kabul edilen varsayımların her sektör için geçerli olmadığını ortaya koymaktadır.

Araştırmanın ikinci bölümünde, ithalat bağımlılığının Kamçı Etkisi üzerindeki rolü incelenmiştir. Uzun tedarik sürelerinin talep dalgalanmalarını artırdığı düşünülse de analizler tedarik süresinin tek başına Kamçı Etkisi'ni artırmadığını göstermektedir. Bunun yerine, bilgi paylaşımının eksikliği, tedarik zinciri esnekliğinin yetersizliği ve yanlış talep sinyallerinin daha önemli olduğu belirlenmiştir. Bu bulgular, şirketlerin yalnızca tedarik sürelerini kısaltmaya odaklanmak yerine, tedarik zinciri iş birliğini ve bilgi paylaşımını artırmaları gerektiğini göstermektedir.

Tezde ayrıca, COVID-19 pandemisinin Kamçı Etkisi üzerindeki etkileri endüstri düzeyinde analiz edilmiştir. Pandemi, küresel tedarik zincirlerinde ciddi aksaklıklara neden olmuş ve talep dalgalanmalarını artırmıştır. Pandemi öncesi ve pandemi dönemi verileri karşılaştırıldığında, talep belirsizliği yüksek olan sektörlerin Kamçı Etkisi'nden daha fazla etkilendiği görülmüştür. Bu sonuçlar, kriz dönemlerinde tedarik zincirlerinin dayanıklılığını artırmak için risk yönetim stratejilerinin uygulanmasının önemini ortaya koymaktadır.

Bu zorlukları ele almak amacıyla, çalışmanın son bölümünde Kamçı Etkisi'ni azaltmaya yönelik gelişmiş tahmin yöntemleri incelenmiştir. Geleneksel tahmin yöntemleri, doğrusal olmayan talep desenlerini yakalamakta yetersiz kalmakta ve yanlış talep sinyallerine neden olarak Kamçı Etkisi'ni artırmaktadır. Bu tez, makine öğrenmesi tabanlı tahmin yöntemlerinin Kamçı Etkisi'ni azaltmada etkili olduğunu göstermektedir. Özellikle Uzun Kısa Süreli Bellek (LSTM) ve LightGBM modellerinin tahmin doğruluğunu artırarak talep dalgalanmalarını azalttığı ve stok yönetimini iyileştirdiği belirlenmiştir.

Bu tez, Kamçı Etkisi'nin endüstriler üzerindeki etkilerini ampirik olarak inceleyerek, belirleyici faktörlerini ortaya koymakta ve etkili tahmin yöntemleri sunmaktadır. Elde edilen bulgular, geleneksel varsayımların her sektör için geçerli olmadığını göstermektedir. Tedarik süresinin Kamçı Etkisi'ni doğrudan belirleyen bir faktör olmadığı, ancak bilgi paylaşımının ve tedarik zinciri iş birliğinin büyük önem taşıdığı belirlenmiştir. Çalışmada sunulan bulgular, Kamçı Etkisi'nin olumsuz etkilerini azaltmaya yönelik stratejiler geliştirmek isteyen tedarik zinciri yöneticileri için yol gösterici niteliktedir.

1. INTRODUCTION

The well-known Bullwhip Effect (BWE), or demand amplification, is a significant supply chain phenomenon first conceptualized by Forrester (1961) in his systems dynamic simulation analysis. His findings revealed that small fluctuations in customer demand can lead to disproportionate increases in production rates, subsequently resulting in backorders, excess inventory, and capacity underutilization at upstream levels. Forrester's pioneering work provided a foundation for understanding the complex feedback loops inherent in supply chain systems.

Building on these insights, Sterman (1989) introduced the Beer Distribution Game, a role-playing simulation that modeled supply chain decision-making across different tiers—retailer, wholesaler, distributor, and manufacturer. Regardless of participants' expertise or background, the game consistently demonstrated that demand and inventory fluctuations were amplified at each stage, highlighting that the Bullwhip Effect is not merely a result of individual decision errors but rather a structural characteristic of supply chains.

Early analytical research on BWE originated with Kahn (1987), who developed a production decision model under serially correlated demand. His findings suggested that even an optimal production strategy could result in higher variance in production than in actual sales, due to the inherent tendency of firms to adjust production levels based on forecasted rather than actual demand. Lee et al. (1997) extended this analysis by incorporating inventory management, price fluctuations, and order batching, demonstrating that even in the absence of lead time delays, repeated forecasting at different supply chain levels can independently generate BWE.

Following the establishment of theoretical models, empirical studies sought to validate the prevalence and cost implications of BWE across industries. Metters (1997) and Chen and Lee (2012) confirmed its presence in real-world supply chains, while subsequent research by Shan *et al.* (2014), Isaksson and Seifert (2016), and Zhao *et al.* (2019), quantified its financial and operational consequences, such as higher inventory holding costs, increased production variability, and service level inefficiencies.

Despite these well-documented adverse effects, the overall impact of BWE on firm performance remains debated. Studies such as (Mackelprang and Malhotra 2015) and (Baron et al. 2018), suggested that, in some contexts, BWE does not necessarily reduce profitability, as firms may find it more cost-effective to tolerate demand fluctuations rather than invest in mitigation strategies. These findings challenge the traditional notion that BWE is inherently detrimental, highlighting the need for a nuanced, industry-specific approach to understanding and addressing demand amplification.

Quantifying the true economic impact of BWE poses two major challenges. First, defining the scope of adverse effects is inherently complex, as some associated costs (e.g., demurrage fees due to excess warehousing) may not be directly captured in traditional inventory cost models. Such indirect consequences often materialize in aggregate financial metrics like gross margin or total supply chain cost, making causal attribution difficult. Second, isolating inefficiencies directly caused by BWE remains problematic. Standard supply chain performance indicators are influenced by multiple decision-making factors and external shocks, leading to confounding effects that obscure the precise contribution of demand amplification.

To address these challenges, this dissertation examines BWE at an industry level, systematically investigating its drivers, impacts, and mitigation strategies. This research is structured around four key studies:

1. Industry-Level Effects of BWE: This study evaluates how inventory days, backorders, overtime hours, gross job loss rates, and capacity utilization rates correlate with BWE across different industries. While theoretical research suggests that inventory accumulation and backorders should exacerbate BWE, empirical evidence at the industry level remains limited. By analyzing industry-specific data, this study provides a refined understanding of the real-world implications of BWE and identifies sectors that exhibit natural mitigation

mechanisms such as production smoothing.

2. Impact of Importation on BWE: Global supply chains increasingly rely on imported intermediate inputs, yet the relationship between import dependency and BWE remains underexplored. This study investigates whether longer lead times and reliance on foreign suppliers contribute to greater demand amplification. Importation also influences information flow and financial transactions, as firms must navigate customs regulations, supplier coordination, and payment mechanisms that impact demand variability and risk management strategies.
3. The Role of COVID-19 in Amplifying BWE: The pandemic created unprecedented disruptions in global supply chains, exposing vulnerabilities in demand forecasting, supply responsiveness, and inventory management. This study analyzes pre-pandemic vs. pandemic-era data to assess the extent to which COVID-19 intensified BWE across industries. By identifying the sectors most affected, the study offers recommendations for improving supply chain resilience against future crises.
4. Mitigating BWE through Advanced Forecasting Techniques: Recognizing the persistent nature of BWE despite advancements in ERP systems and digital collaboration tools, this research explores machine learning-based forecasting methods, including Long Short-Term Memory (LSTM) and LightGBM models. These techniques are tested against traditional time-series models in an order-up-to inventory control framework, with a focus on variance reduction, order fulfillment rates, and inventory stability.

This dissertation introduces new datasets and variables that provide a more comprehensive representation of both the adverse effects and contributing factors of BWE. The methodological framework is designed to bridge gaps in empirical research, particularly by integrating high-granularity firm-level data with broader industry-wide panel datasets.

By leveraging advanced statistical modeling, trend analysis, and forecasting simulations, this research contributes to the ongoing academic debate on BWE prevalence, its impact on supply chain performance, and the effectiveness of modern mitigation strategies. The findings offer practical implications for supply chain professionals, policymakers, and researchers seeking data-driven approaches to manage demand amplification.

Ultimately, this dissertation aims to enhance supply chain resilience by providing empirical evidence, industry-specific insights, and actionable forecasting solutions to combat the persistent challenges posed by the Bullwhip Effect.

The remaining sections of this thesis are structured as follows. Section 2 presents a comprehensive literature review on the Bullwhip Effect (BWE), detailing its theoretical foundations, adverse impacts, and empirical studies. This section also explores the causality between key supply chain factors and BWE, covering aspects such as lead time, lead time variance, supply disruptions, inventory management, ordering costs, information sharing, and financial flow mechanisms that influence demand amplification.

Section 3 outlines the methodology employed in this study, beginning with data acquisition and indicator definition, followed by the criteria for eligibility of observations to ensure robust analysis. It further describes data estimation techniques, including inferring production from sales and inventory data, logging, and differencing adjustments, to ensure comparability across industries. The correlation analysis framework is introduced to examine key relationships, alongside trend assessments using Mann-Kendall analysis. Additionally, this section details the forecasting methodology employed to establish a baseline for the COVID-19 period, incorporating models such as SARIMA, Prophet, RNN, and LSTM.

Section 4 introduces the propositions tested in this study, investigating how inventory levels, backorders, labor intensity, capacity utilization, and reliance on imported inputs impact BWE. This section also examines the long-term trend of BWE and the extent to which the COVID-19 pandemic exacerbated demand amplification across industries.

Section 5 presents the empirical results, systematically evaluating each proposition. This includes findings on the relationship between inventory smoothing and BWE, the role of backorders and job losses, the impact of capacity utilization, and the effect of importation reliance. The results further reveal that while some industries experience strong BWE effects, others exhibit mitigating mechanisms such as production smoothing. Additionally, a comparative analysis of forecast model performance and amplification ratios during the COVID-19 period is included to assess industry-level susceptibility to demand shocks.

Section 6 discusses the managerial implications derived from the findings. It highlights that not all industries respond to BWE in the same manner, challenging conventional assumptions about inventory and backorder relationships. The discussion emphasizes best practices from industries that effectively mitigate BWE and recommends focusing on supply chain collaboration and information sharing rather than simply reducing lead times through nearshoring. The persistence of BWE, despite advances in information technology, suggests that current digital tools are insufficient, necessitating the adoption of AI-driven decision-making and collaboration platforms.

Section 7 introduces advanced forecasting methods designed to mitigate BWE, addressing the limitations of conventional forecasting techniques. This section explores machine learning-driven predictive models that enhance demand accuracy and supply chain performance. It covers key forecasting techniques, including lead-time demand calculations, safety stock optimization, and order-up-to-level inventory control, along with the evaluation of their impact on variance reduction, order fulfillment rates, and inventory stability. The results of one-step-ahead vs. multi-step-ahead forecasting models are also compared to determine their effectiveness in reducing demand volatility.

Finally, Section 8 concludes the thesis by summarizing its key contributions, addressing its limitations, and proposing future research directions. It reflects on the practical implications for supply chain resilience and the potential for further advancements in AI-driven forecasting models. The thesis concludes with a list of references and a biographical sketch of the author.

2. LITERATURE REVIEW

This section provides a comprehensive review of the literature on the BWE, starting with its definition, historical origins, and key theoretical contributions. The discussion then extends to the adverse effects of BWE, drawing from empirical findings that link demand amplification to higher operational costs and inefficiencies in planning across industries. Given the definition of BWE, the relationship between production smoothing and order variability is also explored, considering conflicting evidence on whether inventory investments mitigate or exacerbate demand fluctuations.

The second half of the literature review delves into empirical studies that have measured BWE across different contexts, supply chain structures, and industry settings. This includes discussions on data granularity, aggregation effects, and measurement methodologies, which have influenced the interpretation and generalizability of BWE research.

The causality between key supply chain factors and BWE is examined, with a focus on lead time, supply disruptions, ordering policies, information sharing, and financial constraints. The relationships between these contributing factors and BWE are summarized and visually represented in Figure 2.1, which serves as a conceptual framework for subsequent analysis.

2.1 Definition of Bullwhip Effect

Lee et al. (1997) define the BWE as the rapid increase in the volatility of orders as demand is transmitted upstream in a supply chain, from the end customer to the supplier. They assert that the flow of information, limited to sales orders, tends to be distorted, leading to misinterpretations by upstream parties in their inventory and production decisions. This distortion often results in inefficiencies and increased costs as upstream parties overcompensate for perceived demand variability. In another paper, (Lee et al., 1997) define BWE as “the phenomenon where orders to the supplier tend to have larger variance than sales to the buyer.” This definition implies that demand is distorted as it propagates through upper echelons, counteracting the aggregation effect, and causing a magnification of demand fluctuations.

The term "Bullwhip Effect," also known as the "whiplash" or "whipsaw" effect in some industries, was coined by managers at Procter & Gamble while examining the demand for Pampers disposable diapers. They observed greater fluctuations at the manufacturing level, despite assuming stable demand (Lee et al., 1997). This observation highlighted how even products with relatively predictable consumer demand could experience significant variability in order quantities as they moved upstream in the supply chain. Hammond (1994) introduced a real case where increasing weekly demand fluctuations began straining logistics operations. They discussed contributing factors such as promotions, inventory policies of distribution centers, and full truckloads, along with solutions like information sharing and Vendor Managed Inventory (VMI), as well as challenges like uncooperative distributors. These case studies underscore the practical implications of BWE and the importance of strategic interventions to mitigate its effects.

However, the phenomenon itself was discovered earlier by (Forrester 1961). Following Simon (1952) work on the dynamics of production control, Forrester studied BWE through systems dynamic simulation analysis and termed it “demand amplification.” Forrester demonstrated that while the system is in balance, a 10% increase in demand can lead to a 50% fluctuation in the production rate at the factory level, illustrating the significant impact on orders and inventory levels across the supply chain. This foundational work laid the groundwork for understanding the systemic nature of BWE and its potential to disrupt supply chain stability.

Sterman (1989), building on Forrester's work, simulated the BWE using the "Beer Distribution Game" to study the behavioral and structural complexities of decision-making in supply chains. In this game, a demand shock initiated by the consumer causes orders to spiral out of control throughout the supply chain. This effect is considered an overreaction to demand amplification by decision-makers. The amplification and phase lags in ordering increase steadily the further upstream one moves in the supply chain. Thus, the definition of BWE based on the beer game cannot be generalized to mean just increasing variance in a supply chain but rather increasing variance in response to demand shocks. Despite its limitations, the beer game provides valuable insights into the behavioral responses contributing to BWE, highlighting the role of human decision-making and system feedback loops in exacerbating demand variability.

2.2 Adverse Effects of Bullwhip Effect

Whether defined as an overreaction to demand signal processing shocks or an inevitable consequence of supply chain structure, BWE creates artificial demand variance with no added value to final consumers, affecting all planning activities in the supply chain. The detrimental effects of BWE resemble those of forecast errors. Holt *et al.* (1955) reported forecast error costs in manufacturing environments as financial costs, obsolescence costs, holding costs due to excess inventory, lost sales, penalties due to backorders, and increased setup, overtime, hiring, and layoff costs to address fluctuations. These cost items are summarized as outcomes of BWE by Lee et al. (1997), highlighting the broad range of inefficiencies that BWE can introduce into supply chain operations.

Anecdotal evidence of inefficiencies caused by BWE (Lee et al., 1997) (Hammond 1994) and savings estimates from collaborative supply chain initiatives to alleviate its effects (Wood 1993) (Lee et al., 1997) suggest significant improvement opportunities. These studies provide concrete examples of how firms have successfully implemented strategies to mitigate BWE, such as better information sharing and coordinated inventory management practices. However, these case studies suffer from publication bias and do not provide a comprehensive view of the phenomenon (Isaksson and Seifert 2016). The selective reporting of successful interventions may overstate their effectiveness and fail to account for the complexity and variability of supply chain environments.

Conversely, McCullen and Towill (2002) noted that supply chain managers often underestimated or overlooked the adverse effects of BWE, leading to negligence of the phenomenon in some industries. This underestimation can result in a lack of proactive measures to address BWE, exacerbating its negative impacts. To address this sampling problem, researchers conducted numerous empirical analyses with panel data to observe the prevalence of BWE and its anticipated adverse effects over time. These large-scale studies aim to provide a more comprehensive and unbiased understanding of BWE's impact across different sectors and supply chain configurations.

2.3 Production Smoothing

The earliest empirical studies contributing to BWE literature are rooted in discussions about whether inventory has a smoothing or destabilizing effect. Holt *et al.* (1960) suggested that the costs of production, moving production, and deviating from target inventory levels lead planners to practice production smoothing. However, macroeconomic studies often conflict with this rationale, indicating that the relationship between inventory management and production variability is more nuanced.

Building buffer stock helps minimize supposedly convex production costs by allowing larger quantities to be produced with fewer setup costs. As Blinder (1982) suggested, this is a third option to address demand shocks after increasing prices or production rates. It involves investing in inventory to gain more flexibility, which can be a strategic response to demand variability.

Holt *et al.* (1968) measured fluctuations as mean absolute deviations from the average flow value to find empirical evidence of production smoothing in macroeconomic data. They studied the television industry, which had low component diversity, showing no significant increase in fluctuations at different supply chain levels for television sets between 1950 and 1960. However, decomposing the data into seasonal, cyclical, and random components revealed amplified fluctuations in the cyclical and random components, while the seasonal component showed mixed results. This decomposition approach provides a more granular understanding of how different types of variability contribute to overall production smoothing.

Blanchard (1983) compared the variance of production of U.S. car manufacturers and the sales of their dealers between 1966 and 1979 using a similar decomposition method. Results showed that production variance was larger than sales variance in both cyclical and seasonal components, aligning with Holt *et al.* (1968). On the other hand, Mollick (2004) used Japanese automotive industry data (1985-1994) and found that production variance was generally less than sales variance, possibly due to different manufacturing philosophies and the early adoption of lean manufacturing principles in Japan. This cross-cultural comparison highlights how different production and inventory management strategies can influence the manifestation of BWE.

Beason (1993) also reported that sales variance exceeded production variance in nearly all 40 Japanese industries between 1970 and 1985. This divergence from U.S. data could be attributed to differences in labor practices, such as hiring and firing, instead of building buffer stock to address demand shocks. Such variations underscore the importance of considering contextual factors when analyzing BWE.

Blinder *et al.* (1981) observed that the variance of deliveries to retailers exceeded the variance of actual sales in the U.S. retail industry between 1959 and 1980, suggesting that production smoothing (or delivery smoothing) did not apply to retail dynamics. Caplin (1985) attempted to explain this conflict by showing that an (S, s) inventory policy adopted by retailers generated more variance than actual demand. This finding suggests that certain inventory policies may inadvertently exacerbate BWE.

Kahn (1987) proposed a production decision model under serially correlated demand, demonstrating that even optimal production solutions result in more variance than sales.

His model, based on Equation 1, where y is production, z is actual sales, Δa is the increment in inventory to replenish it to the target level showed that variance in production could not be smaller than sales variance unless there was a negative correlation between actual sales in the previous period and expected future sales. This model provides a theoretical basis for understanding how production decisions can amplify demand variability.

$$var(y) = var(z) + var(\Delta a) + 2cov(z_{-1}, \Delta a) \tag{1}$$

Blinder and Maccini (1991) questioned the contradiction between macroeconomic and microeconomic theories over inventory, asking if inventory destabilizes the macroeconomy while stabilizing the microeconomy. They found that production variance always exceeded sales variance across industries (retail, wholesale, manufacturing) between 1959 and 1986, challenging the rationale for production smoothing. This contradiction highlights the complexity of inventory management and its broader economic implications.

Bray and Mendelson (2015) introduced a new course in the discussion by proposing a measure of production smoothing performance apart from variance comparisons. They modeled a production and forecasting environment using sales, inventory, price, and confidence index data related to the automotive industry. Their results showed that BWE and production smoothing could coexist, indicating that traditional variance measures might not fully capture the dynamics of production smoothing. This innovative approach provides new insights into how firms can manage production variability while mitigating BWE.

Baron *et al.* (2018) argued that even if BWE is detected, its hindering effects on supply chain performance are not as significant as theory suggests. They found no evidence that BWE negatively impacts firm profitability, suggesting that the adverse effects of BWE might be overstated. This finding challenges the conventional view of BWE and underscores the need for a nuanced understanding of its impact on supply chain performance.

2.4 Empirical Studies for Bullwhip Effect

One of the first empirical studies directly measuring BWE after (Lee et al., 1997) was conducted by Fransoo and Wouters (2000). They discussed how aggregation in product, demand source, and time intervals affects BWE measurement. Their findings were mixed and counterintuitive. While product-level aggregation reduced BWE, it did not reduce demand source-level BWE, contradicting Caplin (1985). This suggests that aggregation effects may vary depending on the level at which they are applied, highlighting the complexity of accurately measuring BWE.

Baganha and Cohen (1998) used U.S. data compiled by the Bureau of Census (1978-1985) to compare the coefficient of variation across production, manufacturing sales, wholesaler sales, and retailer sales stages. They found that wholesaler sales had the greatest coefficient of variation, suggesting demand smoothing by wholesalers. This finding aligned with Blinder et al. (1981), yet manufacturer sales showed the second smallest coefficient of variation, indicating a different pattern. These results underscore the varying impacts of BWE at different supply chain stages.

Lai (2005) studied Spanish retailer SKU-level data, testing for cost shocks, batching, gaming, behavioral reasons, and demand correlation. He found evidence of BWE primarily caused by order batching. This study highlights the importance of understanding specific drivers of BWE within different contexts and supply chain configurations.

Cachon *et al.* (2007) examined industry-level U.S. data (1992-2006), finding that 40% of manufacturing industries exhibited BWE while 60% smoothed production. Similarly, Shan *et al.* (2014) and Bray and Mendelson (2012) found that approximately 67% and 65% of firms, respectively, exhibited BWE. In Cachon *et al.* (2007) study retail industry data did not amplify demand, whereas wholesalers did, mirroring Baganha and Cohen (1998) findings. They noted that smoothing was more prevalent in industries with strong seasonality. This study provides a comprehensive view of how BWE manifests across different industries and the role of seasonality in moderating its effects.

Bray and Mendelson (2012) revisited the U.S. economy to study BWE using quarterly and firm-level data. Their findings conflicted with Cachon *et al.* (2007) showing that seasonal smoothing did not cancel out uncertainty amplification. They also found that the decreasing trend of BWE began before the 1995 information technology milestone. Their work highlights the importance of temporal factors in understanding the evolution of BWE.

Shan *et al.* (2014) used firm-level quarterly data for Chinese companies, partially agreeing with Cachon *et al.* (2007). They observed that seasonality smoothed production, but the trend of amplification was more evident in China. They also confirmed a positive correlation between inventory days and BWE. This study underscores the need for cross-cultural comparisons to understand the varying impacts of BWE in different economic and manufacturing environments.

The granularity and the sampling span are two critical characteristics of datasets used in empirical research. Granularity refers to the level of detail in the observations, encompassing aspects such as product level, time intervals, and organizational units. The granularity of a dataset determines its ability to capture the nuances of the actual flow within the supply chain at various resolutions. Higher granularity provides more detailed insights but may come at the cost of broader generalizability.

Conversely, the sampling span, or coverage, represents the breadth of the data, defining the extent of the economy or industry segments covered by the analysis. While proprietary data from a single firm offers detailed time series across multiple variables, panel data surveys provide equivalent time series across different independent organizations. This breadth is essential for demonstrating the generality and prevalence of the BWE across diverse contexts.

Unfortunately, current datasets often face trade-offs between granularity and sampling span. Industry-level studies with a macro perspective may lack the detailed granularity of single-firm analyses, which can limit the depth of empirical insights. Conversely, the detailed granularity available in single-firm datasets does not capture the broader industry-level trends. These limitations influence both the breadth and depth of empirical studies, impacting the comprehensiveness of findings on BWE.

Data types employed in empirical research on BWE can be classified into three main categories based on their source. These categories affect the granularity in measurements, the number of variables, and the explanatory power or sampling span of the analyses, as summarized in Table 2.1. Understanding the strengths and limitations of each data type is crucial for interpreting the results and implications of empirical studies on the BWE.

Table 2.1: Data Types in Empirical Research

	Prominent Studies Literature	**Advantages**	**Disadvantages**
Single-Firm Data	Fransoo and Wouters (2000) Lai (2005) Chen and Lee (2012) Duan, Yao, and Huo (2015)	The most granularity is available in terms of periods, items, and organization. More variables and additional data sets are available.	Does not provide insight for generalization. Susceptible to publication bias. Data collection may not be consistent.
Firm-Level Data	Bray and Mendelson (2012) Shan, Yang, Yang, and Zhang (2014) Bray and Mendelson (2015) Mackelprang, A. W., & Malhotra, M. K. (2015). Isaksson, O. H. D., & Seifert, R. W. (2016)	Firm level granularity Financial statements allow profitability analysis. A dyadic relationship can be established.	Quarterly level data. No item granularity.
Industry Level Data	Cachon et al. (2007)	The broadest perspective is available for generalization. Different data sets from governmental organizations are available. Monthly data is available for more extended periods.	Industry-level granularity. No item granularity.

The first category, "Single-firm data," provides the most granularity at the item and temporal levels. This type of data can also establish direct relationships with other supply chain partners. While the granularity and identified relationships with other parties enable researchers to study the mechanisms of the BWE more precisely, they do not permit testing the prevalence and magnitude of the phenomenon across the broader economy. Despite this limitation, single-firm studies facilitate comparative analyses to understand the data aggregation effect on measured BWE due to their higher granularity compared to firm-level and industry-level studies. For example, Fransoo and Wouters (2000), one of the earliest single-firm studies, demonstrated how aggregation in product, demand source, and time intervals affects BWE measurement. Their study found that while aggregation at the product level reduces BWE, it does not do so at the demand source level (retailers). Similarly, Lai (2005) examined data from a Spanish retailer, testing for cost shocks, batching, gaming, behavioral reasons, and demand correlation at both SKU and category levels. He concluded that aggregation does not mask the prevalence of BWE but decreases its statistical significance.

The second category, "Firm-level data," allows researchers to associate results with individual firms. These datasets, primarily obtained from the financial statements of publicly traded firms, offer firm-level granularity and financial performance measurements at the expense of quarterly aggregation. As suggested by Cachon et al. (2007), recent studies have favored firm-level data to gain more granularity. However, this data type suffers from temporal aggregation and lower item granularity than single-firm studies. Moreover, it does not provide a comprehensive picture of all industries, being inherently biased towards publicly traded companies that are generally expected to perform better.

The third category is "Industry-level panel data," which provides the most extensive breadth but less depth than the other data types in terms of organizational and item granularity. In this study, industry-level data is preferred for two main reasons. Firstly, industry-level data offers the opportunity to link various factors with their effects on industries based on their characteristics concerning the contributing factors of BWE. Secondly, industry-level data, classified under the North American Industry Classification System (NAICS), allows for the integration of additional datasets, enriching the analysis. The availability of more variables related to BWE enables the exploration of different relationships between the phenomenon and its underlying factors, providing a broader perspective.

2.5 The Causality Between Factors and the BWE

The causality between factors and the BWE is intricate. A causality model is developed to map how essential factors discussed in literature interact with each other. Figure 2.1 provides an overview of the multitude of driving factors contributing to supplier order variance and the BWE. The predominant driving force feeding the causality we modeled is the extended lead times resulting from longer transit durations, which influence other factors. The literature consensus acknowledges that supply chains with extended lead times are prone to amplify the BWE. Based on this argument, we examine other aspects of importation as either positive or negative driving forces for demand amplification.

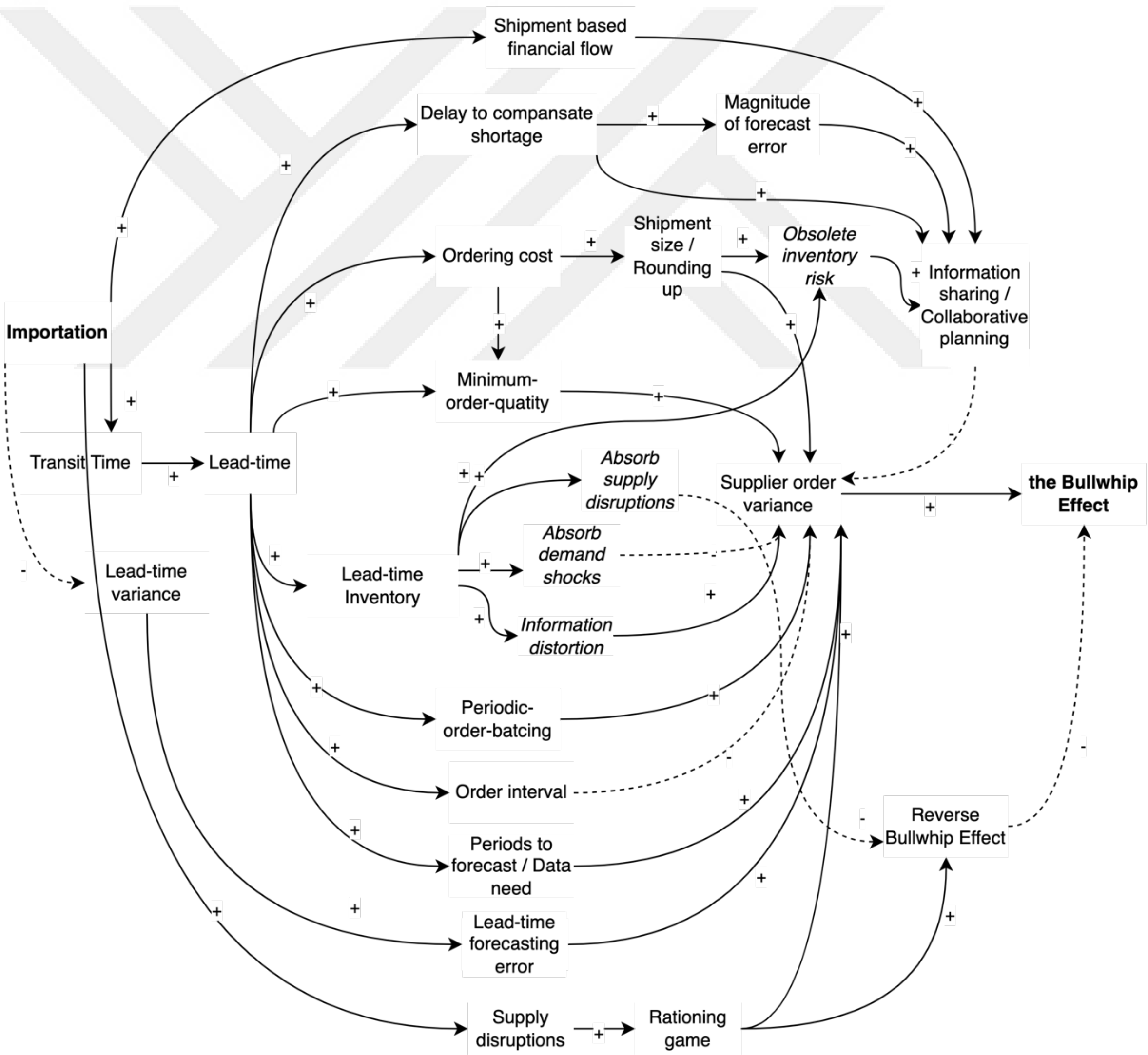

Figure 2.1: Causality Model for Importation and BWE

To identify these driving forces, the following section discusses how these factors amplify or dampen order variance with support from the related literature. Figure 2.1 suggests that most driving factors feed the BWE by increasing supplier variance. Conversely, attenuating factors such as information sharing, reverse BWE, order interval, and lead time inventory compete with amplifying forces.

2.5.1 Lead Time

Lead time in a production and distribution system is the cumulative duration encompassing order processing, production, logistics, and transit time between the supplier and the customer. It plays a critical role in supply chain dynamics and is widely recognized as a contributing factor to BWE.

The positive relationship between the lead time and BWE has been widely accepted by researchers from the beginning of the literature on BWE. In their seminal work, Lee et al. (1997) showed analytically that if the expected demand for subsequent periods is positively correlated with current demand, longer lead times amplify BWE.

However, the involvement of lead time in BWE is not straightforward. In Figure 2.2, we summarized how lead time is interrelated with other contributing factors that end up with two main causes of BWE, as described in Lee et al. (1997). Figure 2.2 highlights how longer lead times often lead to greater order batching intervals, increased supplier order variance, and higher inventory levels, all of which intensify the BWE. Moreover, lead-time forecasting errors can distort demand signals, leading to overcompensation and unnecessary adjustments in production and procurement decisions.

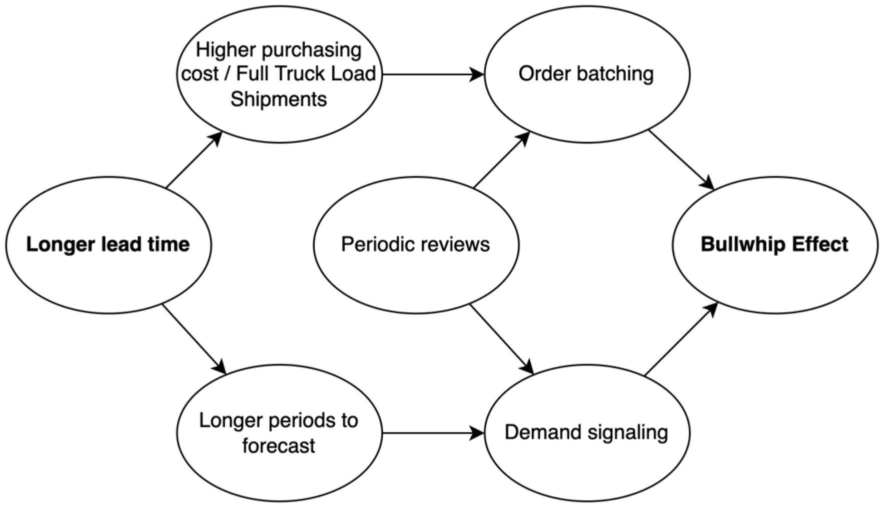

Figure 2.2: Causes and Contributing Factors of BWE in Lee et al. (1997)

Although the longer lead times increase exposure to uncertainty and exacerbate BWE, lead time does not necessarily contribute to BWE (Chen and Lee 2017). If the lead time is covered with deterministic demand or forecast performance is not negatively affected by extended periods, then lead time does not cause BWE (Chen and Lee, 2009). Moreover, some demand processes are less susceptible to amplifying variance over larger lead times Gaalman *et al.* (2022).

2.5.2 Lead Time Variance

Like the lead time, the variance of the lead time is considered only an exacerbating, not a triggering factor for BWE (Fioriolli and Fogliatto 2008, Duc et al. 2008, Heydari et al. 2009, Bolarín et al. 2009, Chaharsooghi and Heydari 2010) unless it is forecasted to determine order quantity (Michna et al. 2020).

As illustrated in Figure 2.1, lead time variance interacts with several factors such as transit time, supplier variability, and shipment delays to influence supply chain fluctuations.

According to Heydari *et al.* (2009) and Chaharsooghi and Heydari (2010), the lead time mean represents the supplier's endogenous performance, and the lead time variance is an exogenous factor. In this context, they showed that the mean of lead time is more significant than the variance regarding BWE, whereas the variance affects total supply chain cost. Nevertheless, the mean and the variance of lead time can be correlated (Fang et al. 2013).

In the case of importation, uncertainties in transit time, customs delays, and congestion contribute to lead time variance, as shown in Figure 2.1. Transporters predict the uncertainty in transit time due to congestion or other import-related delays and inflate the related costs (Taylor et al. 2004). This may imply that forecasted uncertainty may trigger the BWE, as Michna et al. (2020) suggest. However, the status of current global transit times is public knowledge and easy to obtain visibility compared to inventory levels of manufacturers.

Predicting production lead time can be more challenging due to the complexity of product levels, unlike distribution lead time. Production lead time is influenced by various factors, including the manufacturer's process time, inventory position, capacity availability, production lot sizing, and other constraints in a manufacturing environment sensitive to the products being produced. Conversely, distribution lead time relies on more predictable variables, such as transit time, shipment schedules, and customs clearance, independent of the product. Arrival dates are usually certain once the goods have been loaded onto the ships, whereas the same certainty does not apply to sending purchase orders.

Regarding domestic supplies, the supplier's inventory and production capacity play a significant role in determining lead time variance. However, in cases of overseas importing, transit time becomes more significant, and production lead time becomes less so. This relationship suggests that extended lead times in importation due to the logistics component may be more predictable and thus cause less exposure to the BWE.

2.5.3 Supply Disruptions

While supply disruptions are often associated with increased volatility in supply chains, their impact on the BWE can be paradoxical. As illustrated in Figure 2.1, disruptions can either intensify demand fluctuations or, in some cases, lead to a Reverse Bullwhip Effect (RBWE). RBWE, introduced by Rong, Shen, *et al.* (2008), explains that overreacting to random disruptions can increase demand variance, similar to a rationing game. However, it is important to note that the RBWE does not prevent BWE from happening and causes an umbrella shape that impacts the middle of the supply chain most significantly (Rong, Snyder, et al. 2008).

From the importation perspective, exogenous variance caused by random supply disruptions, such as the COVID-19 pandemic, may induce BWE but can be masked by RWBE.

2.5.4 Lead Time Inventory

Longer lead time means more extensive inventory and safety stock, which can be a smoothing or amplifying factor for BWE Xiaojing *et al.* (2008). The safety stock calculations and order-up-to points adjusted to lead time demand inflate the order size and variance. Figure 2.1 illustrates the intricate relationship between lead time, inventory, and demand variability.

However, the role of inventory is not straightforward in the BWE context. Investing in inventory to have more flexibility is an option to address demand shocks, along with increasing the prices or the production rate (Blinder 1982). Therefore, maintaining higher inventory levels can aid in mitigating the amplification of variance as it absorbs the demand shocks and supply disruptions (Chang and Lin 2019). Conversely, in lack of information sharing it can distort the demand signaling for the upstream level (Lee et al. 2000). Since variance ratio measurement is applied to both production smoothing and BWE studies, production smoothing literature sheds some light on the discussion.

In empirical production smoothing literature, there are conflicting results on whether inventory has a smoothing or destabilizing effect. Specifically, in the U.S. perimeter, empirical evidence suggests that inventory can be a destabilizing factor (Holt et al. 1968a, Blanchard 1983, Blinder and Maccini 1991, Beason 1993, Mollick 2004).

Blinder and Maccini (1991), suggested that inventory may stabilize at the firm-level where it destabilizes the industry-level. They discussed that production smoothing may not be prevalent since the observed inventory levels are built up due to lot sizes rather than smoothing purposes. According to this reasoning, since lot sizes are dependent on ordering costs, lead time related inventory can be considered an amplifying factor.

However, Bray and Mendelson (2015) refuted the dichotomy that production smoothing is the opposite of BWE as they showed that BWE and production smoothing can co-exist.

2.5.5 Periods To Forecast (Data Need)

The relationship between forecasting horizon and lead time plays a crucial role in determining the accuracy of demand predictions and, consequently, the magnitude of the BWE. As illustrated in Figure 2.1, longer lead times necessitate extended forecasting periods, increasing exposure to forecasting errors and demand uncertainties.

When lead times are short, firms can rely on more recent and accurate demand data to guide their inventory replenishment decisions. However, as lead time expands, firms are required to predict demand further into the future, where the likelihood of forecast errors rises significantly (Chen, Drezner, et al. 2000, Chen, Ryan, et al. 2000).

The relationship between forecasting horizon and lead time plays a crucial role in determining the accuracy of demand predictions and, consequently, the magnitude of the BWE. As illustrated in Figure 2.1, longer lead times necessitate extended forecasting periods, increasing exposure to forecasting errors and demand uncertainties.

When lead times are short, firms can rely on more recent and accurate demand data to guide their inventory replenishment decisions. However, as lead time expands, firms are required to predict demand further into the future, where the likelihood of forecast errors rises significantly

2.5.6 Ordering Cost, Order Batching, and Order Interval

The logistics component of lead time transit time is an essential driver of the cost of purchasing. Unless a cheaper mode of transport is preferred, such as the railway rather than the road, which is not always a feasible option, longer transit times mean more expensive freight costs due to fuel consumption.

Since longer lead times imply more ordering costs, it inflates the demand as economies of scale drive the decision-making. For instance, planners prefer full-truck loads (FTL) with larger lot sizes rather than less-than-truckload (LTL) with smaller lots that enable more frequent ordering. This factor can be considered the triggering factor of another contributing factor, the order batching.

Order batching is a well-known cause of the BWE that distorts demand signaling, and it shares lead time as a contributing factor (Lee et al. 1997, Holland and Sodhi 2004). As shown in Figure 2.1, order batching distorts demand signals by introducing artificial demand surges at specific intervals, amplifying variance along the supply chain.

The economy of scale encourages order batching to have bulkier deliveries to reduce unit costs. Goods are prone to be batched in larger sizes for long-haul shipments. Therefore, our proposition that tests the effect of lead time on BWE concurrently tests the demand signal processing and order batching as an inducing factor.

However, it is worth noting that there are three distinct implications of order batching. In Lee et al. (1997), the synchronization of orders of multiple retailers is considered order batching as the first implication. The second involves consolidating future period needs due to periodic inventory monitoring. The third entails rounding up orders to meet a multiple of a fixed order quantity.

In his research, Cachon (1999) investigated the effects of various factors on order batching in a similar model to Lee et al. (1997), a 1-supplier, N-retailers supply chain. He examined the impact of the size of retailers' batches, the length of their ordering intervals, and their degree of concurrence. He showed that increased order interval has an attenuating effect on supplier variance. Our product-level analysis also confirmed this argument empirically.

For fixed order quantities, on the other hand, Potter and Disney (2006) suggested that quantities close to average demand are causing the least BWE.

From the importation perspective, transport costs are higher than domestic costs for the exact distances due to the opportunity cost of the uncertainty of waiting times at border crossings (Taylor et al. 2004). Moreover, the study conducted by Hummels and Schaur (2013), showed that the transportation cost for intermediate goods is 60% higher than that of the final goods. Therefore, importation is expected to increase ordering costs.

Nonetheless, the issue of order batching stemming from lead time can be mitigated by implementing alternative logistics strategies. In a recent simulation study by Benrqya and Jabbouri (2022), the researchers employed a model representing a three-tier supply chain. Their findings demonstrated that using a cross-docking approach effectively reduces the BWE by mitigating the impact of order batching. Significantly, this reduction in BWE was observed regardless of the length of the lead time. Thus, as depicted in Figure 2.1, ordering costs, order batching, and order intervals are interconnected mechanisms that influence demand variability and supply chain stability. While economies of scale encourage bulk orders, alternative logistics strategies can mitigate BWE without sacrificing cost efficiency.

2.5.7 Information Sharing and Collaboration

Information sharing is the immediate cure for BWE proposed by Lee et al. (1997), and how information sharing reduces the adverse effects of BWE is extensively studied. It helps upstream parties not to create unintentional shock waves with misinterpreted demand signals from downstream partners. The relationship between information sharing and lead time is also well-studied in the BWE context. The stochastic nature of lead time has an exacerbating effect that can be alleviated by information sharing (Chatfield et al. 2004, Kim et al. 2006, Agrawal et al. 2009).

The importance of information sharing is even more pronounced in longer lead time cases such as importation, where supply chains span multiple regions and countries. The extended nature of global supply chains necessitates greater visibility to mitigate risks associated with longer transit times, customs delays, and supplier reliability issues. As a result, firms engaged in international trade have a stronger incentive to implement collaborative information-sharing practices, as longer lead times inherently increase supply chain risk and volatility.

Based on the literature showing that information sharing is an alleviating factor for lead time related BWE, we presume that longer lead times create an initiative to employ Information sharing that may also reduce BWE as a side-effect.

Moreover, the lack of trust is a barrier to implementing successful information sharing (Kembro et al. 2017). On the other hand, importation depends on trust and strong relationships with global partners. Consistent partnership promotes trust and helps build enduring partnerships, making it easier to address challenges to initiate information sharing, which is as essential as the mutual benefit (Pohlen and Goldsby 2003).

As shown in Figure 2.1, information sharing serves as a mitigating factor that counterbalances lead time-related demand fluctuations. While longer lead times may initially exacerbate the BWE, they also create an opportunity to implement collaborative forecasting and planning solutions, which can ultimately reduce demand variability and stabilize supply chains.

2.5.8 Shipment-Based Financial Flow

Another significant but not quantifiable driver is the shipment-based financial flow between parties. As the distance between parties increases the risk and financing, the gap between payment and usage increases (Schmidt-Eisenlohr 2013). Moreover, enforcing contracts in foreign trade in case of a non-conforming shipment or payment dues is inherently riskier than domestic trade.

In importation, financial flow is critical in managing the risks and uncertainties associated with international trade. This, in turn, influences the dynamics of information flow by fostering trust, cooperation, and adaptability in supply chain relationships and payment methods. Since shipment information is an essential part of the process, shipment-based financial flow improves the timing of the information sharing.

There are three methods for payment transactions in foreign trade: cash-in-advance, where the importer pays before the delivery; bank-intermediated, where payments take place after confirmations; and open account, where the importer is allowed to have the shipment before payment for a specific limit. Although these methods position risks in different handover points, none of them eliminates the risk entirely for all parties. Bank needs to audit the risk of their corresponding firm, which is in another country, increasing the cost of financial flow (International Monetary Fund 2009). Therefore, to manage these risks, companies narrow down the trusted suppliers and sustain their relationships.

According to International Monetary Fund (2009), by 2009, 40% of transactions were open accounts in foreign trade, which is in a declining trend.

As depicted in Figure 2.1, shipment-based financial flow interacts with order fulfillment, trust-building mechanisms, and supply chain transparency, reinforcing the importance of financial stability in mitigating lead time-related uncertainties and BWE.

3. METHODOLOGY

This study employs a comprehensive methodological framework to analyze the BWE across industries. The methodology consists of multiple sequential steps, categorized into four main phases: data acquisition, indicator definition, theoretical alignment, and empirical analysis, including correlation and trend analyses. Each phase is designed to ensure the robustness, validity, and reliability of the findings, with particular emphasis on mitigating potential biases and ensuring comparability across industries.

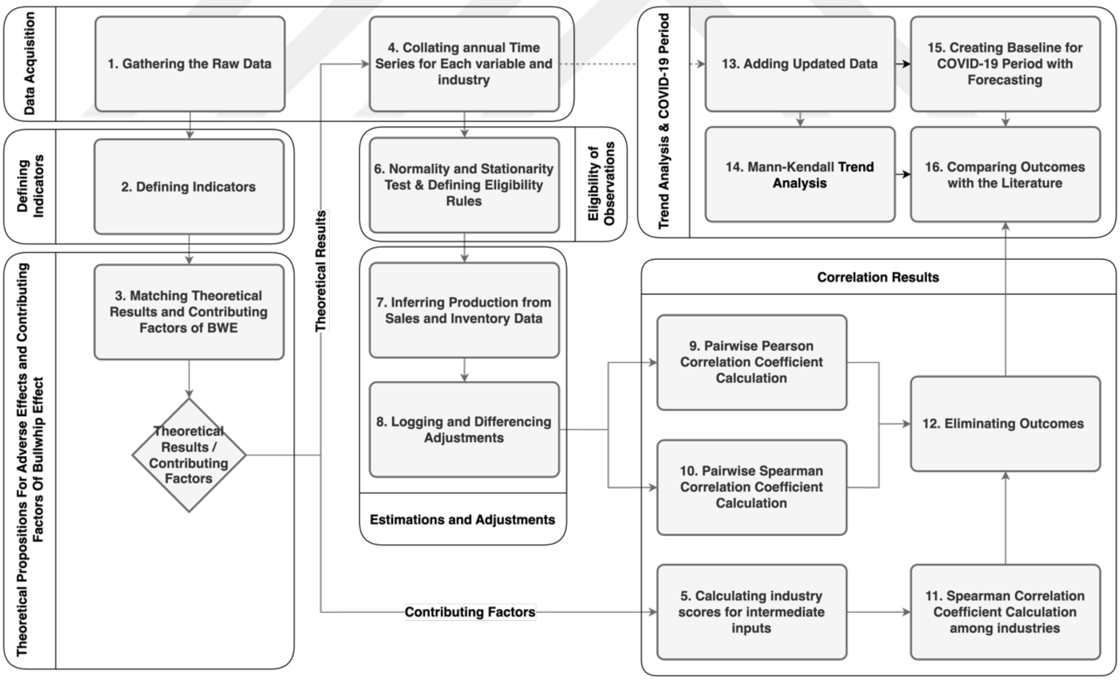

Figure 3.1: Flowchart of Proposed Methodology

Figure 3.1 schematizes the proposed methodology in a flowchart. The initial step in the methodology is data acquisition, which involves obtaining data sets from authorized sources. This step is crucial because data is collected from various institutions, each sharing data in different table formats and with varying coverage. Therefore, a key part of this step is aligning these disparate data formats into a uniform format to ensure seamless communication and integration between them. This alignment is essential for subsequent analysis and comparison.

Following data acquisition, the next step is indicator definition, where specific indicators are defined to measure supply chain management performance. These indicators are carefully selected to match theoretical results and contributing factors of the BWE. For example, the amplification ratio, which serves as an indicator for BWE, and labor intensity, used for complementary analysis, are introduced. The backorder ratio and inventory days are derived from raw data sets, while other indicators like overtime hours, job loss, and capacity utilization are used as they are published.

After defining the indicators, the methodology proceeds with annual compilation, where all time series data are compiled annually at the industry level for variables representing the theoretical results of the BWE. This annual aggregation provides a consistent temporal resolution for comparing different industries and variables.

The subsequent step involves industry ranking, where industry scores are ranked relative to each other based on contributing factors. This ranking helps to identify industries that are more or less susceptible to the BWE based on specific characteristics.

To ensure the validity of the analysis, the methodology includes normality and stationarity testing. This step tests the normality and stationarity of all time series to avoid spurious correlations. An elimination procedure is defined to remove any non-normal or non-stationary series, ensuring that the correlation analyses are conducted on reliable data.

Since direct production data is unavailable, the methodology includes a step for production data inference. In this step, production data is inferred from sales and inventory data. This inference is necessary to calculate key metrics, such as the amplification ratio, which requires both sales and production figures.

For any series that are found to be non-stationary, a non-stationary series processing step is performed to ensure stationarity. This typically involves techniques such as differencing or log-differencing to remove trends or seasonality. For instance, inventory days are differenced to detrend and ensure the stationarity of the series, using log-differencing to adjust for percentage changes between observations.

With stationary data, the methodology proceeds to correlation analyses, where correlation analyses are conducted to investigate the propositions of the literature. These analyses aim to identify relationships between the amplification ratio and various contributing factors.

Following the correlation analyses, a statistical significance testing step is performed. In this step, the statistical significance of the results is tested for all pairs of variables. Non-eligible observations are eliminated to ensure that only statistically significant relationships are considered.

To account for temporal dynamics, the methodology incorporates trend analysis, employing the Mann-Kendall test, which is preferred over linear regression due to its robustness against non-normal distributions. The Mann-Kendall test identifies trends in the BWE over time, with significantly positive Kendall statistic values indicating an upward trend and significantly negative values indicating a downward trend. This rank-based approach is particularly effective in analyses involving outliers and non-normal distributions.

The methodology also incorporates a structured approach to assessing the impact of COVID-19 on supply chains, recognizing the pandemic's significant disruptions. This analysis involves data preparation, forecasting, error analysis, and baseline creation to evaluate whether the BWE intensified during the pandemic. Various forecasting models, including SARIMA, Prophet, RNN, and LSTM, are employed to estimate expected trends, while comparative analyses determine deviations from pre-pandemic conditions. A baseline scenario is constructed to isolate the pandemic's influence, and model performance is assessed by comparing pre-COVID-19 (2016–2019) and COVID-19 (2020–2023) periods. The application of advanced machine learning models, such as Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, ensures a robust evaluation of nonlinear dynamics. Figure 3.5 provides a detailed schematic representation of these methodological steps.

The methodology culminates in reporting and discussion, where the results of the correlation analyses are reported and discussed for all industries, considering their specific characteristics. The divergence among industries is expected to yield a more nuanced understanding of the BWE and the measured performances for further research. This final step synthesizes the findings and provides a nuanced understanding of how the BWE manifests differently across various industries.

3.1 Data Acquisition

Thanks to the data tables provided by Cachon *et al.* (2007), which are digitally published by the authors, most of the calculated values did not need to be reproduced. Some values were partially derived from the raw data tables to modify the structure to fit the temporally disaggregated table format.

Although more recent periods were available, the period between 1992 and 2005 was selected to preserve the continuity of analyses. Since raw data sets are backwardly adjusted by data sources for consistency over the years, extending the time series would create difficulties in aligning new results with those referenced in the study by Cachon *et al.* (2007).

Moreover, extending the period would increase the number of non-stationary time series, jeopardizing the validity of correlation analysis. Five new time series are introduced in this study as new variables, along with two new time series derived from the sales and inventory series, to investigate their association with the BWE. The sources, periods, and number of individual industries are summarized in Table 3.1.

Table 3.1: Data Sets Used in The Study

Data	Source	Period	Interval	Number of Industries
Overtime Hours	Bureau of Labor Statistics	Monthly	1992-2005	29
Gross Job Loss Rate	Bureau of Labor Statistics	Quarterly	1992-2005	17
Capacity Utilization Rate	Federal Reserve System	Monthly	1992-2005	20
Intermediate Inputs	Bureau of Economic Analysis	Irregular	2007	49
Imported Intermediate Inputs	Bureau of Economic Analysis	Irregular	2007	49
Unfilled Orders	U.S. Census Bureau	Monthly	1992-2005	50

Inventory	U.S. Census Bureau	Monthly	1992-2005	86
Shipments	U.S. Census Bureau	Monthly	1992-2005	86
Full-Time Equivalent Employees	Bureau of Economic Analysis	Annual	1997	21
Value Added	Bureau of Economic Analysis	Annual	1997	21

Due to data availability limitations, the scope of the analyses is narrowed to manufacturing industries. Backorders and capacity utilization series are included, at the expense of excluding retailers and wholesalers. This exclusion precludes the investigation of different levels of the supply chain. However, the addition of new variables is preferred over the breadth of analysis to enhance the explanatory power of the study.

All data sets used in this study are obtained from U.S. governmental organizations: the Census Bureau, Bureau of Labor Statistics (BLS), Bureau of Economic Analysis (BEA), and the Federal Reserve.

The Census Bureau and the Bureau of Economic Analysis (BEA) are integral parts of the U.S. Department of Commerce. They provide essential national, regional, industrial, and international accounts, including crucial information on manufacturing, gross domestic product, input-output accounts, and price indexes. The core data sets on sales and inventory are sourced from the Manufacturers' Shipments, Inventories, and Orders (M3) survey provided by the Census Bureau. Intermediate inputs are derived from detailed input-output statistics, indicating the interrelationships between different industries. Additionally, price indexes are employed as auxiliary data to adjust all monetary time series.

The Bureau of Labor Statistics (BLS) is a governmental organization that supports both public and private decision-making by tracking labor market activity, working conditions, price changes, and productivity within the United States. In this study, the gross job loss series is obtained from the Business Employment Dynamics (BDM) program, and the overtime series is sourced from the Current Employment Statistics (CES) program of BLS. As this study considers overtime hours of productive labor, the overtime hours of supervisory employees, which are not typically adjusted to production levels, are excluded. Fortunately, the BLS provides the average overtime hours of nonsupervisory employees separately.

The gross job loss rate indicates the net decrease in employment on a quarterly basis. It comprises the total of establishment closure rates and contraction rates, defined as percentages of the total employment in the same quarter.

The Federal Reserve, the central bank of the United States, publishes an Industrial Production and Capacity Utilization (G.17) survey monthly. This survey provides an index of industrial production and capacity utilization rates across various industrial sectors. The capacity utilization rate for manufacturing industries is obtained from this survey. Capacity utilization rates consider a facility's maximum production rate within a practical work schedule, excluding material availability shortfalls. The Federal Reserve also accounts for capacity reductions due to expected downtimes.

The values for Overtime, Gross Job Loss Rate, Capacity Utilization Rate, Inventory Days, and Back-order to Shipment Rate within each year are averaged to ensure comparability across different time series.

3.2 Defining Indicators

Indicators are measurements to be used as a variable in the correlation analyses. Table 3.2 shows that the backorder ratio and inventory days are derived from raw data sets; overtime hours, job loss, and capacity utilization are used as they are published.

Table 3.2: Indicators for Electronic Equipment, Appliances, and Components

	Amplification Ratio	Overtime Hours	Inventory Days	Inventory Days (Ln Diff'ed)	Capacity Utilization Rate	Gross Job Loss Rate	Backorder to Shipment
1992	1,04	3,18	96		82,22	3,9	1,85
1993	1,27	3,67	88	-0,09	86,58	3,98	1,71
1994	1,02	4,23	84	-0,04	91,76	3,63	1,69
1995	0,9	3,52	87	0,03	91,3	4,05	1,74
1996	0,66	3,67	81	-0,07	89,84	3,9	1,62
1997	0,48	3,89	78	-0,04	88,31	3,98	1,64
1998	0,35	3,54	74	-0,05	85,69	3,9	1,6
1999	0,6	3,65	73	-0,03	83,14	4,05	1,65
2000	0,44	3,72	71	-0,02	85,88	3,78	1,74
2001	1,18	3	76	0,07	76,86	6,28	1,72
2002	0,27	3,06	78	0,02	73,36	5,33	1,74

2003	0,23	3,44	74	-0,05	74,45	4,33	1,72
2004	0,49	4,01	70	-0,06	77,1	3,4	1,74
2005	0,8	3,8	68	-0,02	80,92	3,35	1,8

Two sample data tables show all indicators for one industry (Table 3.2), and one time series for all industries (Table 3.3) is shared for Electronic Equipment, Appliances, and Components industry and inventory days indicator. The same data structure is used for all variables and industries to conduct correlation analyses throughout the study.

Table 3.3: Inventory days for all industries

Industry	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
All with Unfilled Orders	108	102	97	96	95	90	90	88	87	92	86	84	77	76
Aluminum and Nonferrous Metal Products	66	67	64	61	64	59	57	59	61	63	57	60	52	50
Apparel	77	88	91	95	86	79	85	88	91	93	88	86	78	85
Basic Chemicals	83	83	78	78	80	76	80	80	81	84	78	76	73	72
Beverage and Tobacco Products	110	119	107	101	102	108	109	103	101	97	111	110	108	108
Beverage	66	65	64	65	69	77	82	85	87	91	94	94	93	95
Communications Equipment, Defense	67	99	121	118	113	157	237	204	160	151	195	229	187	144
Communications Equipment, Nondefense	126	124	118	126	112	98	105	91	94	115	138	133	121	112
Communications Equipment, Total	116	121	118	125	112	101	110	95	96	116	142	140	126	115
Computer and Electronic Products	123	115	107	102	100	91	92	86	88	106	105	94	83	76
Computers and Related Products	90	90	82	74	68	56	50	45	45	54	52	44	35	31
Construction Machinery	124	108	97	98	95	87	85	95	88	108	105	84	69	65
Construction Supplies	77	72	70	71	69	66	65	63	65	66	63	65	61	63
Consumer Durable Goods	46	43	40	42	43	39	37	34	38	39	35	35	35	37
Consumer Goods, Total	53	52	51	51	51	49	50	48	49	49	47	46	44	43
Consumer Nondurable Goods	55	56	55	55	54	53	55	54	53	53	52	51	48	46
Dairy Product	25	25	26	25	27	26	25	25	27	25	26	27	25	25
Defense Aircraft and Parts	256	253	235	216	187	178	225	224	245	184	152	136	132	146
Defense Capital Goods	179	176	189	173	169	147	157	181	193	166	135	110	108	110
Durable Excluding Defense	88	83	78	78	78	74	74	71	72	77	72	69	64	65
Durable Excluding Transportation	99	94	89	89	88	83	82	81	82	89	84	82	75	75
Durable Goods Total	95	89	84	83	82	78	78	75	76	81	75	72	67	68
Electric Lighting Equipment	107	99	91	95	87	85	80	79	83	83	80	83	80	82
Electrical Equipment	105	94	89	89	83	81	77	76	77	86	86	80	78	75

Electromedical, Measuring, and Control Instrument	156	147	135	133	133	135	133	132	131	138	130	123	104	107
Electronic Computer	93	96	81	72	68	51	42	32	30	38	32	24	19	15
Electronic Equipment, Appliances, and Components	96	88	84	87	81	78	74	73	71	76	78	74	70	68
Fabricated Metal Products	93	88	85	87	84	80	79	78	78	79	77	80	75	78
Ferrous Metal Foundries	72	63	56	57	59	58	62	61	61	68	66	65	57	60
Food Products	39	38	39	39	39	38	37	38	39	39	38	36	34	33
Furniture and Related Products	79	78	78	78	73	70	66	65	67	68	57	62	60	58
Grain and Oilseed Milling	39	36	38	37	41	38	38	37	38	35	34	34	40	44
Household Appliance	74	68	66	74	67	65	62	60	59	54	54	46	47	49
Industrial Machinery	140	140	133	131	131	127	138	130	118	157	150	138	124	142
Information Technology Industries	118	113	107	105	102	93	91	85	85	97	100	94	83	77
Iron and Steel Mills and Ferroalloy and Steel Products	91	89	80	83	86	83	89	92	96	103	91	90	74	89
Leather and Allied Products	97	91	94	107	101	96	105	106	108	132	110	87	86	92
Machinery	117	110	106	107	108	103	102	104	102	113	105	97	90	90
Excluding Defense	74	71	68	69	68	66	66	64	64	67	63	62	58	57
Excluding Transportation	77	75	72	73	72	69	69	68	68	71	68	66	62	61
Material Handling Equipment	102	89	90	90	86	85	82	84	84	97	95	85	77	75
Meat, Poultry, and Seafood Product Processing	18	18	19	19	20	19	20	19	20	19	20	18	17	16
Metalworking Machinery	139	129	125	124	132	124	121	128	120	148	126	114	93	80
Mining, Oil and Gas Field Machinery	241	210	208	202	181	166	193	206	201	177	201	202	171	157
Motor Vehicle Bodies, Trailers, and Parts	48	43	42	43	42	41	39	37	39	39	35	37	37	39
Motor Vehicles and Parts	32	28	26	27	27	26	24	22	25	25	22	23	23	25
Nondefense Aircraft and Parts	222	249	271	265	286	236	222	217	196	176	196	206	181	145
Nondefense Capital Goods	131	126	119	116	115	109	109	105	101	111	109	105	95	90
Nondefense Capital Goods Excluding Aircraft	117	111	105	104	102	94	93	90	89	101	98	93	84	82
Nondurable Goods Total	59	59	57	58	57	56	57	56	56	56	54	53	50	49
Other Durable Goods	85	80	79	79	77	74	72	71	74	75	69	72	68	67
Other Electronic Component	99	87	84	92	93	87	86	78	86	114	100	86	80	68
Paint, Coating, and Adhesive	74	70	70	73	71	70	70	69	68	70	63	59	57	60
Paper Products	62	61	57	54	60	57	57	55	54	54	54	55	49	50
Paperboard Container	59	57	54	59	56	55	54	52	53	50	49	49	48	49
Pesticide, Fertilizer, and Other Agricultural Chemical	74	75	67	69	70	71	74	91	94	99	83	71	75	70
Petroleum and Coal	34	35	34	34	32	32	34	30	26	27	28	32	29	27

Products														
Petroleum Refineries	33	34	33	33	31	31	34	29	25	26	27	31	29	26
Pharmaceutical and Medicine	132	141	143	138	140	135	139	142	149	144	134	131	142	138
Photographic Equipment	97	108	99	107	117	107	110	106	107	103	120	115	108	115
Plastics and Rubber Products	64	61	59	63	61	60	59	57	59	58	53	56	54	57
Primary Metals	77	76	70	70	72	69	71	72	75	80	72	72	62	67
Printing	53	48	49	53	48	45	44	42	43	43	41	43	45	48
Pulp, Paper, and Paperboard Mills	61	63	57	49	62	59	57	54	53	53	56	57	51	52
Search and Navigation Equipment, Defense	200	168	172	167	179	158	147	150	147	147	127	100	95	82
Search and Navigation Equipment Nondefense	312	288	293	260	295	269	270	266	256	253	272	266	238	233
Ships and Boats, Total	84	82	77	82	84	75	65	66	68	83	65	59	60	58
Textile Products	75	75	76	79	77	76	77	72	76	75	68	63	59	59
Textiles	60	61	60	63	62	61	61	63	64	70	58	59	53	53
Tobacco	242	301	250	218	210	210	193	159	150	129	166	162	155	150
Total Capital Goods	139	133	128	123	121	113	114	112	108	117	112	106	97	92
Total	78	75	72	71	71	68	69	67	67	69	65	63	59	59
Transportation Equipment	83	76	69	66	67	63	65	61	62	60	53	50	48	50
Turbines, Generators, & Other Power Trans. Eq'ment	116	114	108	108	102	93	89	88	87	87	77	68	70	75
Ventilation, Heating, A.C., and Refrigeration Equipment	87	76	76	79	76	77	67	69	73	76	68	63	64	65
Audio and Video Equipment	50	45	43	53	62	53	48	49	57	63	53	49	52	56
Automobile	17	15	12	13	13	13	11	10	10	10	10	10	12	12
Battery	85	71	69	73	70	66	65	57	64	67	64	58	58	56
Computer Storage Device	96	78	87	75	86	72	61	69	72	86	92	102	87	97
Farm Machinery and Equipment	93	84	82	82	86	74	82	102	93	85	81	76	65	70
Heavy Duty Truck	34	32	28	28	28	26	24	22	28	35	28	29	28	29
Light Truck and Utility Vehicle	13	10	8	9	9	8	8	7	9	9	8	8	8	10
Miscellaneous Products	128	119	120	119	121	117	114	113	114	113	105	106	106	102
Nonmetallic Mineral Products	84	78	70	73	69	67	64	62	64	67	64	68	65	61
Other Computer Peripheral Equipment	80	82	85	86	59	59	63	66	70	74	87	73	64	66
Wood Products	54	53	54	53	53	50	50	50	54	55	53	51	46	48

Although they are not used to represent the theoretical results or contributing factors of BWE, the amplification ratio, which is the indicator for BWE, and labor intensity which is used as complementary analysis, are introduced in this section.

Table 3.4: Descriptive Statistics for Amplification Ratio and Other Variables

	Amplification Ratio	Inventory Days	Backorders to shipment	Capacity Utilization Rate	Overtime	Gross Job Loss
N	1204	1204	700	280	406	238
Min	0,014	6,8	0,0	60,7	1,8	2,8
Max	98,9	311,5	30,8	95,3	9,4	11,7
Range	98,9	304,8	30,7	34,6	7,6	8,9
Median	0,7	75,1	2,2	80,9	4,6	4,9
Mean	1,1	83,6	4,0	80,7	4,8	5,1
Var	9,9	2242,1	25,9	31,6	1,7	2,4
Std. Deviation	3,1	47,4	5,1	5,6	1,3	1,5
Coef. Variation	2,8	0,6	1,3	0,1	0,3	0,3

Table 3.4 summarizes descriptive statistics for all variables between 1992-2005 without segmenting industries. Including the amplification ratio, seven different time series are added to the data set. The data availability for industries is not the same for all industries. The number of observations ranges between 238 and 1204, which creates a bottleneck for the count of pairwise comparisons.

The main takeaways can be summarized for distinct indicators. The mean amplification ratio is larger than 1 (1.1), implying the prevalence of variance amplification across the industries. However, the coefficient of variation shows that it strongly varies among industries. Though there is a significant declining trend, the mean Inventory days of industries are 83,6 days. Industries have approximately %19 slackness in their capacity. Finally, the average weekly overtime hours is 4,8 for all industries without weighting the number of employees.

3.2.1 The Bullwhip Effect / Amplification Ratio

A common method for measuring the BWE is by calculating the ratio between the demand variances at upstream and downstream levels (Equation 2). Cachon et al. (2007) emphasized that this ratio can be viewed as the balance point between amplifying and attenuating forces, where values greater than 1 represent BWE and values less than 1 represent production smoothing. Therefore, this measurement indicates the strength of variance-amplifying forces relative to production-smoothing forces. For a binary classification of the presence of BWE, the amplification ratio should be greater than 1.

$$Amplification\ ratio = \frac{V_{Production}}{V_{Demand}} \quad (2)$$

Figure 3.2a illustrates how the amplification ratio is calculated using the sales and production data of the same echelon. This method does not require a dyadic relationship between two parties, allowing the amplification ratio to be measured independently of whether the entity is a retailer (R), wholesaler (W), or manufacturer (M).

Another method for measuring BWE involves comparing the coefficients of variation between neighboring stages or dividing the coefficient of variation of production by that of sales. A two-stage (dyadic) relationship between a manufacturer and a specialized wholesaler may offer a consistent linear supply chain structure, as shown in Figure 3.2b.

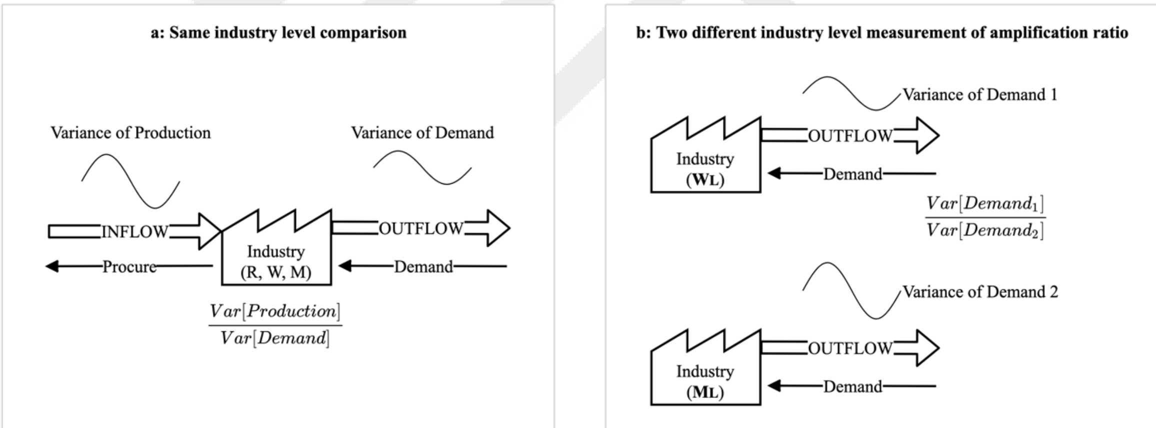

Figure 3.2: Measured flows

Holt *et al.* (1968) discussed the risk of distortion by criticizing the hide-leather-shoe chain proposed by Suits and Mack (1958), where a hide manufacturer may sell leather to both apparel and shoe manufacturers. Holt *et al.* (1968) proposed that the T.V. set industry stages, strictly linked with its exclusive supply chain partners, are suitable for establishing an intra-industry relationship.

However, this approach significantly narrows the applicable industries and reduces the sampling span. Since it is impossible to establish such a relationship across all industries, and to avoid the risk of ambiguity, industry material flow is preferred over a linear supply chain structure.

Also as Chen and Lee (2012) pointed out, representing demand with shipments creates a gap between observed material BWE and information BWE originally defined by Lee et al. (1997). Whereas material flow follows the information flow and approximates the original definition, the amplification ratio based on material flow is subjected to physical constraints.

In the context of our study, we acknowledge that logistics operations act as a limiting factor, similar to production capacity. Material BWE may be masking the information BWE. However, the risk of underestimating does not lead to the false detection of BWE, jeopardize the study's accuracy that shows bullwhipping industries.

Notably, Chen *et al.* (2017), showed that longer lead times decrease the discrepancy between information-based and material-based BWE.

3.2.2 Inventory days

This conventional inventory performance metric measures the time required to deplete inventory, expressed in days (Cox et al. 2004). It reflects the investment in inventory relative to the establishment's sales, normalizing the data and facilitating comparisons across different industries by converting the unit of measurement from currency to days. The price and margin-adjusted shipment and inventory data from the study by Cachon et al. (2007) are utilized to calculate this indicator. The formulation for annual measurement is provided in Equation 3:

$$Inventory\ days\ (day) = \frac{The\ average\ value\ of\ inventory\ at\ standard\ cost}{Annual\ cost\ of\ goods\ sold\ (COGS)\ /\ 365} \tag{3}$$

3.2.3 Backorders to Shipments

To analyze backorders (unfilled orders) and compare different industries, demand series are divided by the corresponding value of shipment data (Equation 4). Since derivative data is a ratio, no further adjustment is necessary.

$$Back\ orders\ to\ Shipments\ (\%) = \frac{Value\ of\ unfilled\ orders\ in\ month}{Value\ of\ total\ shipment\ in\ month} \qquad (4)$$

Backorders do not necessarily indicate stock-outs or lost sales. The Census survey captures unfilled orders at the end of a period. A company may have the inventory for the demand but might delay the shipment due to other causes such as order processing time, order-batching, or logistical bottlenecks. However, it can serve as a good proxy for stock-outs for finished goods and waiting lines for supply.

3.2.4 Labor Intensity

Although not included in any correlation analysis, labor intensity is considered to enhance our understanding of how vulnerable the industry is to BWE as it becomes more labor-intensive. The conventional value-added criterion is used to measure the labor intensity of industries, defined as value-added per employee (Equation 5). In this method, industries are ranked, and those with less than the average value added per employee are considered labor-intensive.

$$\text{Value added per employee} = \frac{\text{Value Added}}{\text{Full Time Equivalent Employees}} \qquad (5)$$

Since the scope of this study concerns the capacity decision mechanisms of individual industries, the labor intensity of intermediate industries is not considered, contrary to the suggestion by Lary (1968).

3.2.5 Imported Intermediate Inputs

We propose intermediate inputs to find an empirical ground for our study. Intermediate inputs refer to the goods and services purchased from all sources, including energy, raw materials, semi-finished products, and services, that are utilized in the production process to create other goods or services rather than for final consumption. It can be calculated by subtracting value added from gross output Bureau of Economic Analysis (2006).

Despite the significance of the importation and the lead time, they have not been given sufficient attention in empirical studies. This lack of attention can be attributed to the unavailability of industry-level data for these variables.

Nevertheless, some approaches have been developed to overcome this limitation. In their analysis of company-level data, Rumyantsev and Netessine (2007) employed publicly available financial statements and used accounts payable outstanding as a proxy for lead time. They assumed that payment terms correlated with shipment lead time.

In contrast to their domestically sourced counterparts, imported goods often encounter delays attributable to many factors, including customs clearance, tests, regulations, and payment transactions. These delays inherently extend the time required for imported intermediates to traverse the supply chain. Therefore, the ratio of imported intermediates represents a good approximation to longer lead times to supply. Similar to our study, Lee *et al.* (2022), utilized import sourcing as a substitute for longer lead times. Using panel data from Korean firms, they empirically showed that longer distances and higher transport costs increase the BWE. By establishing a connection between lead times and the ratio of imported intermediates, our research findings provide valuable insights into the correlation between lead time and the BWE.

The data is derived from the supply-use framework of the Industry Economic Accounts published by BEA, which covers U.S. manufacturing industries. Although the data table is prepared according to NAICS codes, the values are given at the commodity level rather than the industry level. To align the commodity and industry level tables, the 405 commodity groups are matched with the 49 groups of industries, which also have amplification ratio data availability (Table 3.5). Energy and service items are reduced from the lists to include only raw materials and semi-finished products flow.

Table 3.5: Data Sets for Ratio of Imported Intermediate Inputs

Data	Source	Period	Interval	Number of Industries
Intermediate Inputs	Bureau of Economic Analysis	Irregular	2007	49
Imported Intermediate Inputs	Bureau of Economic Analysis	Irregular	2007	49

In terms of the periodicity of the report, BEA published the data only in 2007 and 2012. As the 2007 data is closer to the period we studied, it was used instead of the 2012 data. However, it should be noted that the ratios of most industries did not change significantly over the years when comparing the 2007 and 2012 data.

We propose a ratio of imported intermediate inputs as a variable representing the presence of importation in the corresponding industry.

$$\text{Imported Intermediate Inputs (\%)} = \frac{\text{Value of imported intermediate goods}}{\text{Value of all intermediate goods}} \tag{6}$$

Equation 6 demonstrates the metric as the percentage of imported intermediate goods out of the total intermediate inputs for the corresponding industry. Since the definition covers only intermediate inputs, imported finished goods to be resold by the importer are excluded from the analysis.

3.3 Eligibility of Observations

All correlated time series are also annually grouped and averaged at the industry level. Consequently, the number of samples for all correlation coefficient calculations was 14 (the number of years between 1992 and 2005). For a sample size of 14, the significance levels above 20% are excluded from the analyses.

Due to the limited sample size for annual industry-based comparisons, the Shapiro-Wilk test was conducted to inspect the normality of these annual sums. Among 288 industry-based annual sums, 229 indicated normality. If a time series failed to indicate normality, its Pearson coefficient was neglected.

While most series indicated normality, the Spearman method, a non-parametric rank-based correlation, was also employed to avoid questionable coefficients Kowalski (1972). Thus, an observation is eligible if:

1. There is statistical significance for the Spearman coefficient, regardless of the normality test.
2. The Spearman coefficient is not statistically significant, but the Pearson coefficient is, and both the amplification ratio and tested series pass the Shapiro-Wilk test.

All time series grouped by industry and variable are tested for stationarity to avoid spurious correlations. Given the small sample size, the annual Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test is preferred over the Dickey-Fuller test to detect trend stationarity. While most variables (251 out of 288, except for inventory days) were stationary, a significant number of non-stationary series were found for inventory days (41 out of 86). To address this, inventory days are log-differenced to ensure stationarity. Other time series are not log-differenced to maintain sample sizes. The few non-stationary series of amplification ratios are excluded to ensure consistency in analysis.

3.4 Estimations and Adjustments

The fundamental estimations and adjustments from the study by Cachon et al. (2007) are presented to demonstrate how production data is inferred from inventory changes and how time series are logged and differenced, as applied to new data sets.

3.4.1 Inferring Production from Sales and Inventory Data

Production data derived from inventory changes is formulated in Equation 7, suggesting that if inventory decreases, the industry produces to fulfill the demand minus the net change in inventory Blinder *et al.* (1981). Therefore, the negative inventory change and demand are produced during that period. Conversely, if the inventory level increases, the industry invests in inventory to fulfill the demand and increase the inventory. This formulation implies that the industry invests in inventory with the sum of demand and net inventory change in the exact period. Cachon et al. (2007) also define production in the same fashion.

$$Y_{it} = S_{it} + (I_{it} - I_{it-1}) \tag{7}$$

$Y_t = X_t + \Delta N_t$; "where Y is real output for manufacturers or deliveries for wholesalers and retailers, X is real shipments, and ΔN is real inventory investment" Blinder and Maccini (1991).

The validity of this equation is straightforward for retailers and wholesalers. However, inventory flow for manufacturing industries with raw materials, work-in-progress (WIP), and finished goods inventories is not intuitive and requires clarification.

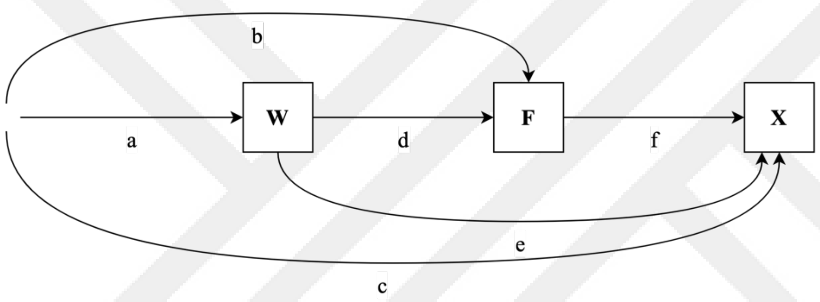

Figure 3.3: Stages[1] of Processing Blinder (1986)

Figure 3.3 and Equation 8 illustrate how inventory consumption for a typical manufacturing process is decomposed into phases , showing that inventory changes equal the difference between sales and production.

$$\Delta F + \Delta W = a + b - (e + f) \tag{8}$$

Where ΔF is the change in the stock of finished goods, ΔW is the change in the stock of work in progress; therefore, $\Delta F + \Delta W$ is equal to $Y - X$, output minus sales.

[1] a: raw materials, b: components, c: direct shipments to customers, d: work-in-process used in finish goods, e: work-in-process shipped to customer, f: finished goods shipped to customer.

3.4.2 Logging and Differencing Adjustments

There has been a steady decreasing trend in inventory days over the years. Of 86 industries (some aggregated and overlapped), only 8 have a positive slope.

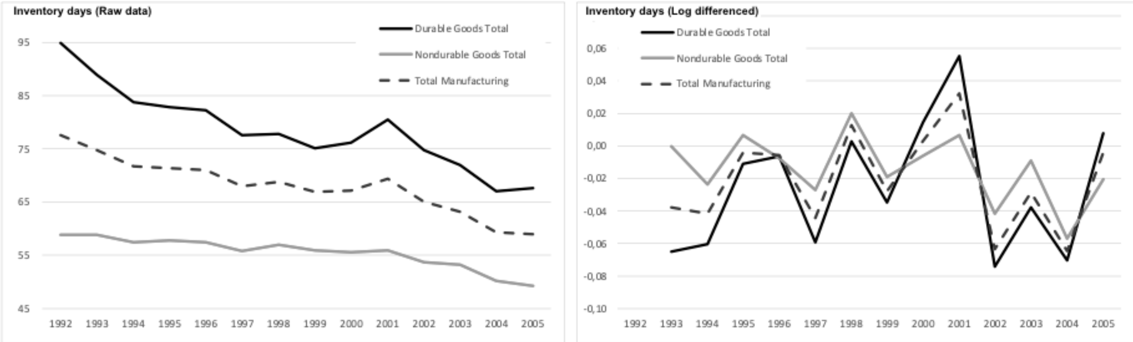

Figure 3.4: Inventory days over the years for unadjusted and log differenced series

Figure 3.4 shows inventory days decrease in aggregated level manufacturing industries. Due to the negative slope among nearly all industries, the data is detrended with log-differencing the time series.

Inventory days are differenced to detrend and ensure stationarity of the series (Equations 9 and 10). The series are logged prior to the first differencing. The adjusted series is the approximate change percentage between observations in the series:

$$\ln(X_i) - \ln(X_{i-1}) \approx (X_i - X_{i-1})/X_{i-1} \tag{9}$$

All index values X_t is replaced by the differenced series,

$$\Delta X_t = X_t - X_{t-1} \tag{10}$$

where X is the average of monthly inventory days values of year t.

Cachon et al. (2007) conducted the Dickey-Fuller test and detected a unit root in production and demand series with 170 monthly observations. To remove it, they log-differenced both series. Since the unit root test has low power for small sample sizes and this study has 14 annual values for each industry, stationarity is investigated with the KPSS test instead of the unit root (Cochrane 1991).

3.5 Correlation Analysis

The theoretical propositions presented in Section 5 were investigated for all industries. Pearson and Spearman correlation coefficients were calculated to explore pairwise associations between variables and the amplification ratio for each industry. Plot boxes are presented for each comparison to provide a general overview of the characteristics of bullwhipping and smoothing industries.

While the Pearson coefficient indicates a linear relationship, the Spearman coefficient indicates a monotonic relationship between variables. Both correlation methods were employed for all pairs of variables, and the strongest coefficient among these methods is highlighted. Nevertheless, for most pairs, both methods yielded similar results.

The amplification ratio was calculated for each year using monthly sales and production time series to obtain more observations. Although quarterly grouping would generate more observations, the quarterly calculated amplification ratio values were too erratic to be used. Conveniently, the annual calculation covers observations for an entire year, eliminating the need for seasonal adjustment.

Since the relationships between variables can vary in direction across different industries, both cases are discussed with their supporting literature.

3.6 Adding Updated Data for Trend Analysis and COVID-19 period

The test data set is split into two periods: pre-COVID-19 (2016 – 2019) and COVID-19 (2020 – 2023). The aim was to identify the models that consistently deliver the most reliable and accurate prediction in the first period. Since the demand and supply shocks during the COVID-19 period can increase the errors as expected, the 2016-2019 period is considered a baseline to assess model performances.

The demand series has been price and margin-adjusted to translate sales figures into deflated cost values. For SARIMA, an Augmented Dickey-Fuller (ADF) test was carried out to determine the stationarity of the demand series. For non-stationary series, differencing and logarithmic conversions, are applied to stabilize variance and mean. Similar transformations are implemented for inventory series to ensure consistency in data. Since Neural Network based forecasting models perform poorly for seasonal and non-stationary demand Zhang and Qi (2005), detrending and deseasonalizing are carried out as key steps in preparing time series data for RNN and LSTM methods along with Minmax normalization. Detrending removes long-term trends to stabilize the series' mean over time, which is important for models that need data to be stationary. The deseasonalizing step removed seasonal patterns, leaving behind the residual component that shows random variations. The original series is decomposed into an additive model's trend, seasonal, and residual components. After making predictions, the trend and seasonal components are added back to the predictions to make them interpretable and to ensure they reflect the original data's scale and behavior.

3.7 Creating Baseline for COVID-10 Period

Figure 3.5 presents a comprehensive overview of this study's forecasting methodologies and BWE analysis. The process begins with collecting historical industry-level data on inventory and demand. To prepare for analysis, these series undergo several preprocessing steps, such as deflation, margin adjustments, deseasonalization, detrending, and normalization. Subsequently, four distinct forecasting techniques, SARIMA, Prophet, RNN, and LSTM, were used to predict a forecasting horizon encompassing the pandemic and the subsequent recovery phase periods.

Once the forecasts are generated for 2020 to 2023, they are integrated into a 'Forecast pool,' where they undergo error analysis. In parallel, observed data is processed similarly to create a comparative basis. Both forecasted and actual data are used to calculate production and, subsequently, amplification ratios, which serve as a metric to assess the magnitude of the BWE in the industry. By comparing these ratios, this study aims to determine whether the pandemic has exacerbated the BWE, indicating a greater variance in inventory and demand fluctuations than typically observed.

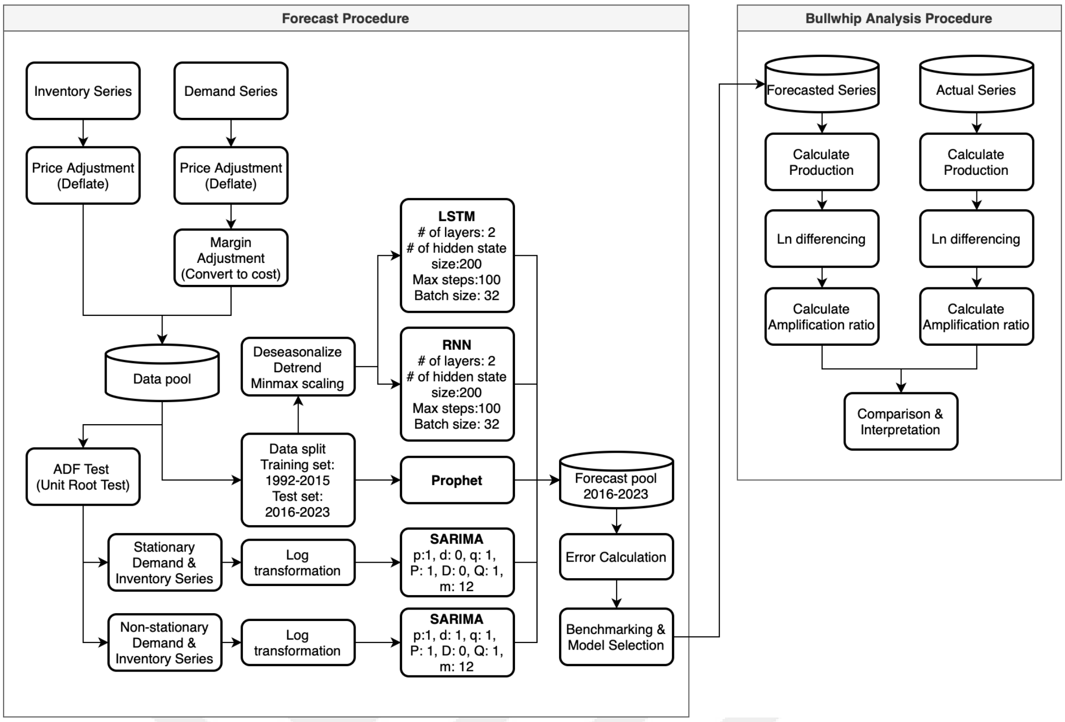

Figure 3.5: Flowchart of Procedures

3.7.1 Forecasting Models

Since BWE exists regardless of COVID-19, distinguishing between endogenous BWEs and those induced by external shocks of COVID-19 remains challenging. To tackle this, we employ advanced forecasting techniques to establish a baseline for the COVID-19 period, which represents what the scenario might have looked like without the pandemic's influence. This baseline is then compared with actual data to evaluate the vulnerability of specific industries to the BWE under pandemic conditions.

In the forecasting phase, Demand and Inventory series for 77 industries are forecasted with SARIMA, Prophet, RNN, and LSTM methods for COVID-19 impacted periods. Advanced models like Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks are applied to address nonlinear dynamics that might be present in the data. This multi-model approach ensures a comprehensive analysis by leveraging the unique strengths of each forecasting technique, which is crucial for strategic planning in the context of unpredictable impacts such as those caused by a global pandemic.

SARIMA, is notable for its ability to explicitly incorporate seasonality through interpretable parameters, making it a benchmark for subsequent models. It is denoted as SARIMA$(p, d, q)(P, D,Q)m$, where p is the autoregressive, d is the differencing, and q is the moving average order, plus *(P,D,Q)* representing the seasonal autoregressive, seasonal differencing and seasonal moving average order, respectively. Finally, m is the number of periods in each season. The SARIMA equation is represented in Equation 11.

$$(1-\sum_{i=1}^{p}\phi_i L^i)(1-\sum_{i=1}^{P}\Phi_i L^{ixm})(1-L)^d(1-L^m)^D X_t = (1+\sum_{i=1}^{q}\theta_i L^i)(1+\sum_{i=1}^{Q}\Theta_i L^{ixm})\varepsilon_t \tag{11}$$

where

- L is the lag operator, $L^k X_t = X_{t-k}$
- ϕ_i, Φ_i are the coefficients for the autoregressive and seasonal autoregressive terms.
- θ_i, Θ_i are the coefficients for the moving average and seasonal moving average terms.
- X_t is the time series and ε_t is the error term at time t

Prophet is a forecasting tool developed by Facebook®, designed to produce high-quality forecasts for time series data that exhibit patterns on different time scales. It is particularly well-suited for data with strong seasonal effects and several seasons of historical data. In addition to SARIMA, Prophet can automatically detect trends, adjust the forecast to reflect growth that is not necessarily steady, and flexibly accommodate seasonality. The core formula of Prophet is given as Equation 12.

$$y_{\mathrm{t}} = g_{\mathrm{t}} + s_{\mathrm{t}} + h_{\mathrm{t}} + \epsilon_t \tag{12}$$

where,

– y_{t} is the predicted value at time t,
– g_{t} is the trend component, modeling non-periodic changes over time,
– s_{t} is the seasonality component, capturing periodic changes,
– h_{t} is the effects of holidays or special events,
– ϵ_t is the error term at time t.

Recurrent Neural Network (RNN) is an artificial neural network that recognizes patterns in data sequences. Unlike traditional neural networks that process inputs independently and assume inputs and outputs independence, RNNs have loops that allow information to persist, which is suited for tasks where context and temporal dynamics are crucial, like in time series forecasting. RNN involves an input layer where the data sequence is received, a hidden layer where the input is processed by combining it with the current state of the network, producing a new state, and an output layer that generates the next value of the time series. The fundamental formula for a basic RNN involves a simple update rule for the hidden state (Equation 13).

$$h_t = \sigma(W_h h_{t-1} + W_x x_t + b) \tag{13}$$

where,

- h_t is the hidden state at *t* $and h_{t-1}$ is the hidden state at *t-1*
- x_t is the input at *t*
- W_h andW_x are weights matrices for the hidden state and input, respectively.
- b is the bias vector
- σ is the activation function

Long Short-Term Memory (LSTM) networks are a special kind of RNN models, capable of learning long-term dependencies in data sequences. Introduced by Hochreiter and Schmidhuber (1997), LSTMs are specifically designed to avoid the long-term dependency problem typical of standard RNNs, making them highly effective for a wide range of sequence prediction problems where context from the distant past is informative for predicting future events. LSTM networks introduce a more complex update mechanism that involves several gates to control the memory and flow of information:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{14}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{15}$$

$$\tilde{C}_t = tanh(W_C[h_{t-1}, x_t] + b_C) \tag{16}$$

$$C_t = f_t(C_{t-1}) + i_t(\tilde{C}_t) \tag{17}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{18}$$

$$h_t = o_t \tanh(C_t) \tag{19}$$

where,

- f_t, i_t, o_t are forget, input, and output gates outputs, respectively.
- $\tilde{C}_t$ is the new cell candidate vector,
- C_t is the cell state vector,
- h_t is the output hidden state vector,
- $W\ and\ b$ represent weight matrices and bias vectors for different gates,
- $\sigma\ and\ tanh$ are activation functions,

This study employs a systematic approach to forecasting demand and inventory series from 1992 to 2023. The methodologies summarized in Figure 3.5, integrated into our analysis, include traditional and advanced forecasting models. Monthly data from 1992 to 2016 was used to train forecasting models, and 2016 - 2023 data was used to evaluate their performance and confirm their predictions. The accuracy of forecasts is assessed using traditional non-scale-based error metrics, such as mean absolute percentage error (MAPE), which provides insights into each model's predictive accuracy and reliability, facilitating objective benchmarking and model selection process. Subsequently, for the Bullwhip Analysis, the amplification ratio for forecasted and actual values between 2020 and 2023 is calculated.

3.8 Mann-Kendall Trend Analysis

Contrary to Cachon *et al.* (2007), we preferred the Mann-Kendall test to carry out trend analysis as a better option than linear regression analyses, which require normality.

Unlike linear regression, which assumes normality, the Mann-Kendall test is a non-parametric method used to detect trends in time series data without requiring a specific distribution. It evaluates whether a monotonic trend (increasing or decreasing) exists by analyzing the relative ordering of data points rather than their actual values. The test computes the Kendall statistic, which counts the number of concordant and discordant pairs in the dataset. A significantly positive Kendall statistic value indicates an upward trend, while a significantly negative Kendall statistic value suggests a downward trend. By using this rank-based approach, the Mann-Kendall test is more robust against outliers and non-normal distributions, making it a preferable choice for trend detection in this analysis.

4. PROPOSITIONS

The study by Cachon et al. (2007) proposed hypotheses to investigate the relationship between the amplification ratio and theoretical contributing or mitigating effects. They hypothesized that the amplification ratio negatively correlates with seasonality, price variations, and demand shocks. While the authors confirmed all hypotheses except the effect of demand shocks on the BWE, which remained inconclusive, this study aims to test the associations between the amplification ratio and widely recognized symptoms of the presence of the BWE. BWE can impact the supply chain in several ways (Bhattacharya and Bandyopadhyay, 2011), Xiaojing *et al.* (2008). Figure 4.1 summarizes how these adverse effects align with the included data sets.

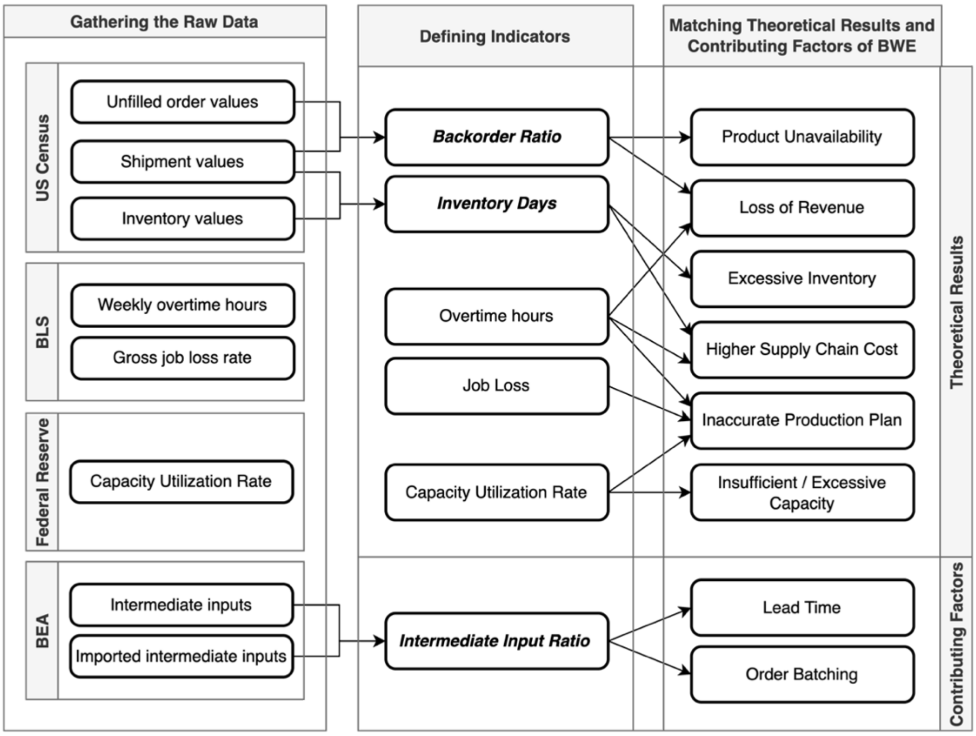

Figure 4.1: Data, Indicators, and Implications

The first three symptoms are measured with straightforward proxies, whereas the last three are indirectly associated with several measurements. For instance, inventory days are used as a proxy for excess inventory, and capacity utilization values represent insufficient or excessive capacity. Product unavailability can be directly associated with unfilled orders, even if the demanding party is willing to accept the lead time.

Overtime hours (correction cost), backorders (unfilled orders, possible lost sales), and higher inventory days (holding cost) can approximate higher total supply chain costs. As noted by Cachon et al. (2007), it is impossible to determine the underlying cause of these symptoms solely from amplification ratio measurements in industry-level data. However, it is feasible to correlate the relative strength of amplification against smoothing with these symptoms to establish an empirical foundation.

These propositions aim to explore how different industries are affected by the BWE and contributing factors rather than to entirely falsify or prove a relationship between variables. Consequently, these propositions are not addressed with hypothesis testing.

4.1 Proposition 1: Inventory Days and Backorders positively correlated with the Amplification Ratio

Empirical evidence of a positive association between inventory days and the BWE is reported by Shan *et al.* (2014) in their firm-level study of Chinese companies. Shan *et al.* (2014) suggested that high inventory is a symptom of poor-quality inventory management, which yields a high amplification ratio. Moreover, longer inventory cycles and inventory decisions for extended periods make firms more vulnerable to demand information distortions. Another recent study by Zhao *et al.* (2019) also confirms the positive association between inventory days and the amplification ratio.

Cachon et al. (2007) did not investigate the association between inventory days and the amplification ratio. To address this gap, inventory days are calculated using the same shipment and inventory data and then compared with amplification ratios for manufacturing industries.

Increased backorders, or unfilled orders, are considered an adverse effect caused by the BWE. This study considers them a proxy for stock-outs or product unavailability. Both metrics indicate how well inventory management performs. To extend prior discussions on what these metrics indicate, their association with the BWE is investigated.

4.2 Proposition 2: Overtime and Gross Job Loss positively correlated with the Amplification Ratio

Overtime is one of the quickest measures to increase capacity, especially if production is labor-intensive and there is excess capacity in fixed resources. In the context of the BWE or production smoothing, it is a correction cost to attenuate demand shocks and level capacity.

The extensive literature on labor economics discusses short- and long-run labor demand and adjustment costs. As reported by Hamermesh Daniel S (1993), most empirical and estimation model studies show a preference for adjusting working hours over headcount. Chang (1983) for the U.S. automobile manufacturing industry and Nickell (1984) for the U.K. manufacturing industry developed employment models under different economic expectations. Both studies showed that overtime is a preferable adjustment over contraction or layoff.

To the best of our knowledge, this is the first study investigating the association between employees' working conditions and the BWE. This association is significant in two ways. Firstly, it explores how the BWE affects employees' welfare by investigating if the BWE harms employees regarding working conditions. Overtime and gross job loss are considered indicators of the industry's working conditions. Secondly, it tests if the BWE increases overtime costs, as suggested by Lee et al. (1997). In their seminal work, they considered overtime one of the many symptoms of the BWE, along with other correction costs.

4.3 Proposition 3: Amplification Ratio and Capacity utilization are positively correlated.

Poor capacity utilization is considered an adverse effect of the BWE. However, the amplification ratio may not increase if the capacity utilization ratio is fixed and saturated. Two relatively recent studies support this argument. Buchmeister *et al.* (2014) reported that higher overall equipment efficiency (OEE) levels, accompanied by efficient inventory management, create less deviation in production and demand, leading to greater stability. Alternatively, as suggested by Chen and Lee (2012), capacity constraints restrain firms from amplifying variance and promote production smoothing.

Long-term investments require upward flexibility if no excess equipment and machinery capacity exists. Thus, these positive upswings may lead to the BWE in capital investment (Anderson et al. 2000). However, in the short term, the BWE is expected to be dampened due to capacity constraints (Chen and Lee 2012). Capacity utilization rate data published by the Federal Reserve is added to the data set to study this interrelation. The hypothesis is defined against this argument to test the widely accepted understanding in the literature.

Importation significantly impacts the material, information, and financial flows that constitute the supply chain. It primarily affects the material flow through its influence on lead times. Transporting goods over long distances, along with customs clearance processes and other potential delays, elongates the time it takes for materials to move from the point of origin to their destination within the supply chain.

4.4 Proposition 4: Reliance on imported inputs positively correlates with the Amplification Ratio

Importation and local supply refer to two different methods of acquiring goods or products, each involving distinct aspects of commerce. Importation involves bringing goods or services into one country from another, with these goods produced or manufactured in a foreign country and transported across borders for consumption or use in the importing country. In contrast, local supply involves the production or provision of goods and services within the domestic borders of a country.

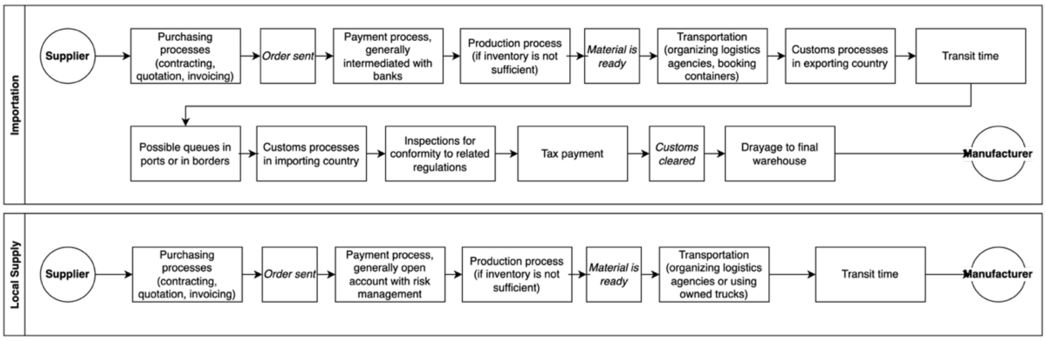

Figure 4.2: Typical Steps of Importation and Local Supply

Figure 4.2 summarizes the steps of both supply structures. Importation involves a more complex supply chain that includes customs clearance, payment of import duties and taxes, adherence to trade regulations, and compliance with foreign regulations. Conversely, local supply chains are typically more straightforward, with fewer steps.

The importation steps and their sequence differ depending on shared borders, modes of transportation, special customs agreements, and payment agreements. Nevertheless, each of these extra steps in importation includes processing times and delay risks, leading to longer lead times and vulnerability to delays for imported supplies.

Based on the data provided by World Integrated Trade Solution (WITS), summarized in Figure 4.3, around 71% of the intermediate goods imported into the U.S. manufacturing industries in 2007 were sourced from countries that do not share borders with the U.S. (The World Bank 2022). Since the data does not define energy, all the flow can be considered seaborne trade.

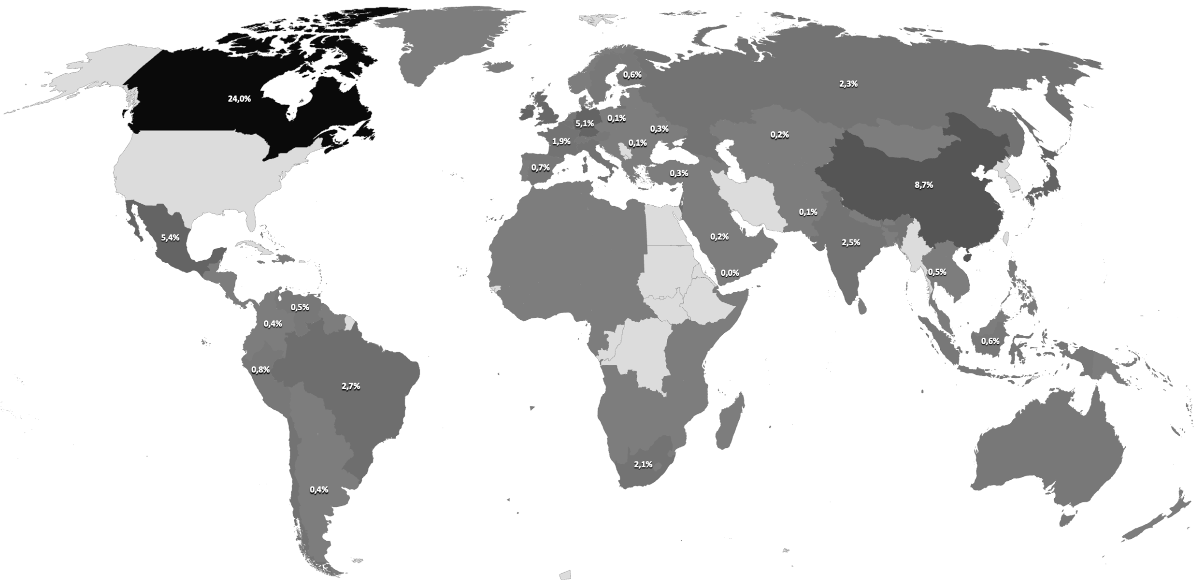

Figure 4.3: Imported Intermediate Goods in 2007

The average transit time for sea freight from China, one of the leading exporting countries, to the U.S. varies between three to six weeks, depending on the coasts, excluding inland transportation. Disney *et al.* (2016) reported that between March and September 2011, the average lead time for transportation between ports in the U.S. and China ranged from three to five weeks. The specific locations of these ports are not detailed in the study. However, according to the U.S. Department of Transportation Maritime Administration, %50,2 of seaborne container trade in U.S. customs ports have arrived on Pacific coasts, which tend to have greater transit time (Maritime Administration 2023).

To compare the transit times of import and domestic trade, we utilized the data from the Freight Analysis Framework of Bureau of Transportation Statistics. According to the data, in 2017, 92% of domestic trade flows, excluding energy commodities, occurred between points less than 750 miles apart, the legal limit for one day's travel time (Bureau of Transportation Statistics 2023).

Therefore, compared to local supplies, importation involves longer lead times in the U.S. due to more delaying steps and far longer transit times. Mexico and Canada, which share a border with the U.S. and constitute 29% of all intermediate inputs, are subject to the same extra importation steps. Thanks to trade agreements such as NAFTA and USCMA, more streamlined processes reduce customs delays. Even though they have shorter transit times than overseas countries, they still face delays and extra costs in the customs process (Taylor et al. 2004); (Avetisyan et al. 2015). Additionally, the prominent mode of transportation for imports from these countries uses slower transportation modes compared to domestic supply. According to data from the Freight Analysis Framework of the Bureau of Transportation Statistics in 2017, 110,000 tons, 114,000 tons, and 115,000 tons of imported products from these countries were shipped via truck, rail, and water, respectively, whereas truck shipments, the fastest mode of transport after air, constituted more than 83% of all domestic flow.

Consequently, most imported goods in the U.S. are subject to longer transit times and delay points compared to local supplies, resulting in extended lead times, a critical factor in the BWE literature.

The act of importation, wherein firms source goods and materials from international suppliers, fundamentally influences the dynamics of the supplier-buyer relationship within a supply chain. This intricate interplay is discussed in Section 1.2. Within this context, competing factors that amplify or attenuate the supplier order variance result in different amplification ratios, which can be empirically traced.

Moreover, firms dependent on imported materials and products often operate with longer lead times. This enables us to establish a link between lead time and the BWE. Therefore, the objective of this study is to gather empirical evidence of the impact of importation and lead time on the BWE.

Through empirical analysis, this study aims to reveal the extent to which importation practices and lead times distinctly contribute to the emergence and mitigation of the BWE, thereby enhancing our understanding of the complexities inherent in supply chain dynamics.

Most factors related to importation feed the supplier order variance. Therefore, the degree of reliance on imported inputs positively correlates with the magnitude of supplier order variance and, consequently, the amplification ratio within the supply chain.

This proposition can reveal the outcome of many attenuating forces in the BWE phenomenon defined in Section 2.1. Conversely, longer lead times to supply may force firms to have better forecasting and planning processes, similar to the effects of seasonality. It can act as a limiting factor that suppresses the BWE, as it does for production capacity.

To address recent discussions on localization as a remedy to alleviate supply chain risks arising from importation.

4.5 Proposition 5: The trend in BWE is decreasing

While advancements in supply chain management practices, widespread adoption of information technologies, and improved demand forecasting techniques have contributed to mitigating the BWE in some industries, the overall trend has remained stable. Cachon et al. (2007) found evidence suggesting a decline in BWE ratios over time, but with variations across industries—some experiencing a decrease while others saw an increase. Additionally, their analysis indicated that BWE was not significantly driven by specific time periods. While a declining trend was observed among retailers and wholesalers, no statistically significant differences were found across time series samples.

Given that our study extends the empirical analysis over a longer period, A trend analysis is conducted to examine whether the decline in BWE persists over time.

4.6 Proposition 6: The COVID-19 pandemic has significantly increased BWE within global supply chains.

The COVID-19 pandemic has significantly amplified the BWE within global supply chains by exacerbating demand fluctuations, inventory imbalances, and supply disruptions. The unprecedented external shocks caused by the pandemic, including abrupt changes in consumer behavior, logistical constraints, and production slowdowns, have intensified variance amplification across industries. Therefore, the pandemic has not only disrupted supply chain stability but has also magnified demand-supply mismatches, reinforcing the presence of an elevated BWE during crisis conditions. By establishing a forecasting-based counterfactual baseline, it is possible to distinguish pandemic-induced distortions from inherent supply chain fluctuations.

This proposition is propelled by a dual motive: to seek empirical evidence of the BWE amidst the pandemic by examining industry-level panel data and to assess the ability of time series analysis to forecast the BWE in terms of amplification ratio.

5. RESULTS

5.1 Proposition 1: Inventory Days and Backorders positively correlated with the Amplification Ratio

Surprisingly, inventory days and backorders, considered significant adverse effects of BWE, did not show a dominant positive correlation with the amplification ratio. Along with a positive correlation that concurs with the findings of Shan *et al.* (2014) Zhao *et al.* (2019), many industries exhibit a negative association between inventory days and variance amplification. Out of 86 industries, the inventory days of 16 are negatively associated with the amplification ratio. Four of these 16 industries indicate BWE in the cumulative calculation, highlighted in black in Figure 5.1. Three of them are aggregated industries. Only six industries show a positive correlation between inventory days and the amplification ratio. Ventilation, Heating, Air-Conditioning, and Refrigeration Equipment Manufacturing is not eligible due to the failed Shapiro-Wilk test.

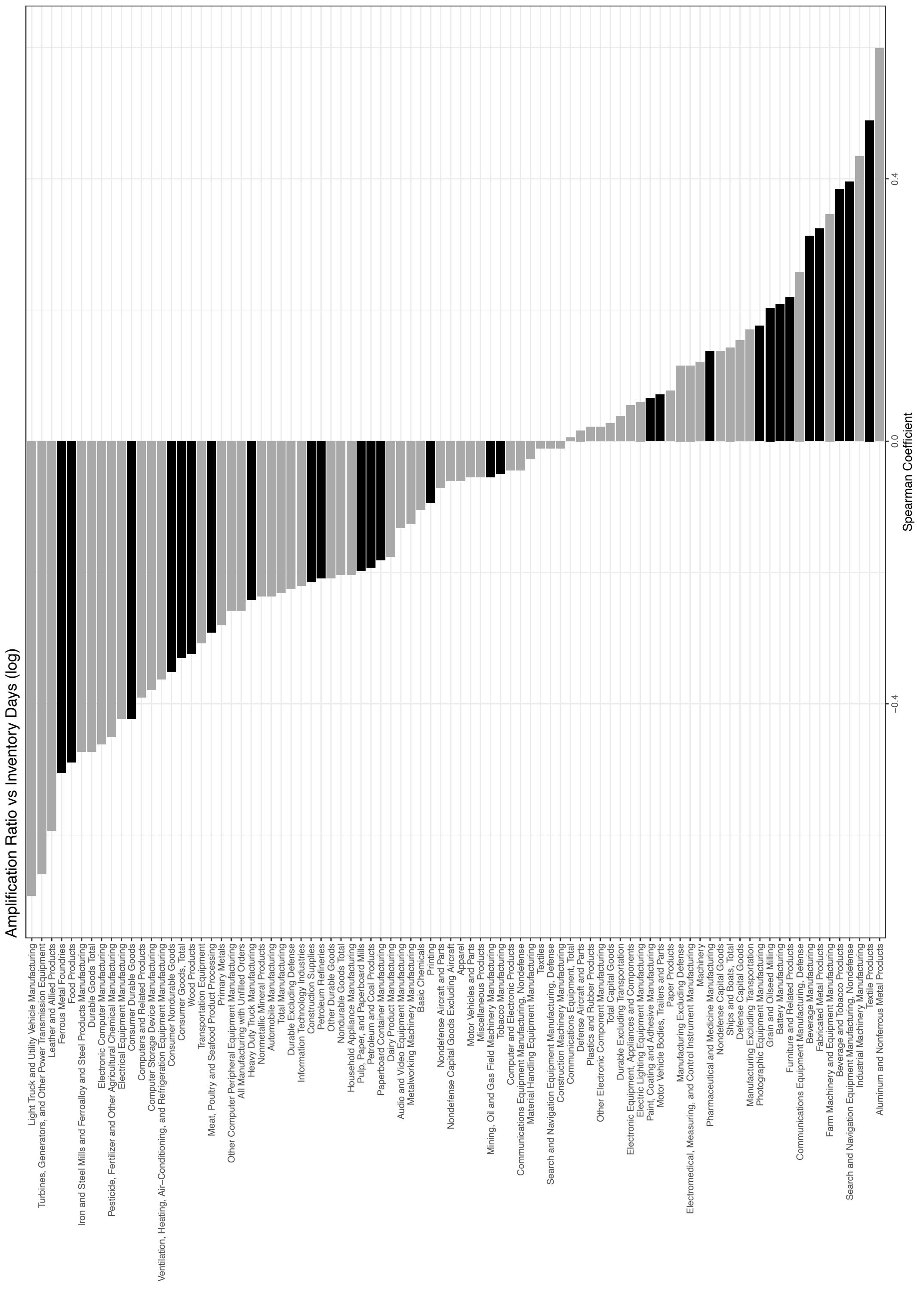

Figure 5.1: Bar-Chart: Amplification Ratio vs. Inventory

Figure 5.2 shows that inventory days are lower in industries exhibiting BWE compared to those that do not. Given that more industries tend to have a negative correlation, this can be interpreted as industries with more inventory days may exhibit production smoothing. Light Truck and Utility Vehicle, Turbines, Generators, Other Power

Transmission Equipment, and Ferrous Metal Foundries, which have low inventory risk due to make-to-order strategies or raw material-dominant inventory, have the strongest negative correlation.

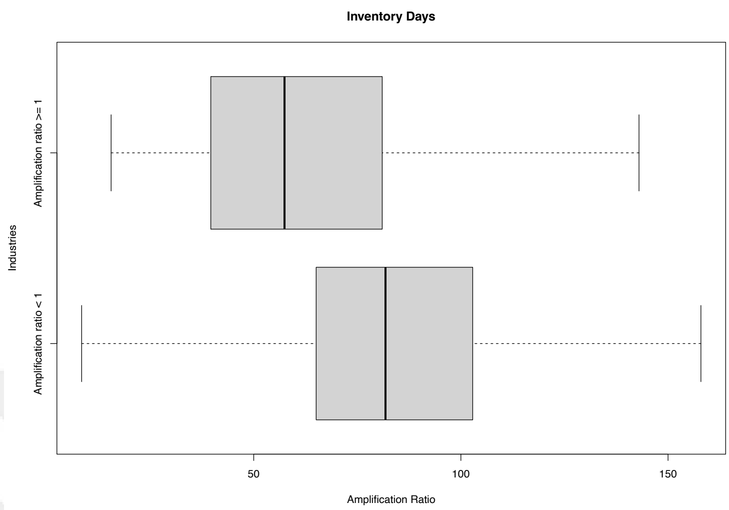

Figure 5.2: Box-Plot: Amplification Ratio vs. Inventory

However, Food Products and Ferrous Metal Foundries have a negative correlation and an amplification ratio greater than one. This result aligns with the study Bray and Mendelson (2015) which argues that the BWE phenomenon and production smoothing can coexist. Results may support the argument that companies may experience BWE even if they practice production smoothing. Textile Products can be affected by high obsolescence risk or seasonality due to fashion trends, negatively affecting higher inventory days. Although Industrial Machinery Manufacturing is a make-to-order industry with less finished product inventory risk, it is sensitive to inventory days in terms of the amplification ratio.

In both cases, these results may be underestimated due to aggregation. Yao *et al.* (2021) suggest that BWE decreases inventory management performance, but temporal aggregation may mask 75% of extra inventory and 25% of stockouts.

Backorders or unfilled orders are considered possible lost sales and indicate uneven demand and supply rates.

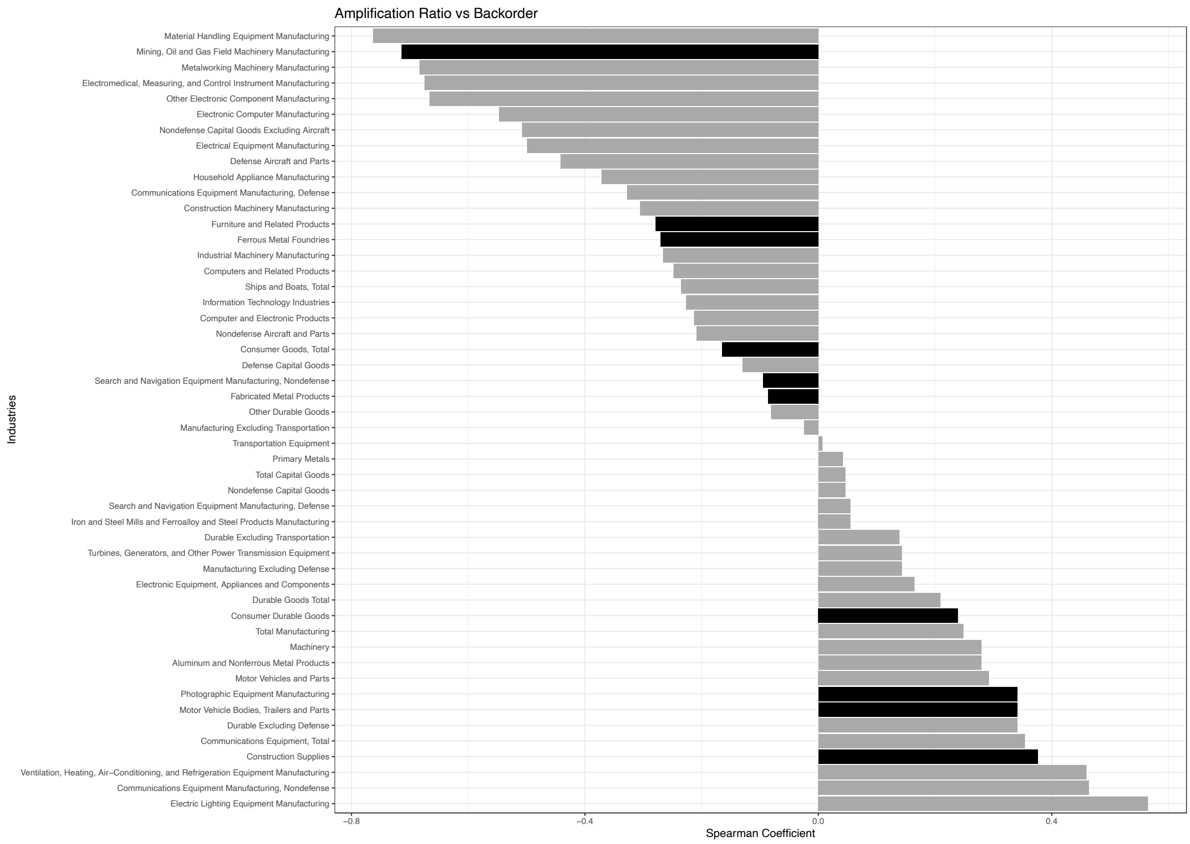

Figure 5.3: Bar-chart: Amplification Ratio vs. Backorder Ratio

Figure 5.3 shows that 12 industries out of 50 have a negative association, whereas five industries are positively correlated. Figure 5.4 demonstrates that the backorder ratio is lower among amplifying industries. The production strategies may justify this. Most of the industries where the amplification ratio negatively correlated with backorders are capital equipment industries or practice make-to-order strategies, such as Nondefense Capital Goods Excluding Aircraft, Mining, Oil and Gas Field Machinery Manufacturing, Metalworking Machinery Manufacturing, Electromedical, Measuring, and Control Instrument Manufacturing, Defense Aircraft and Parts, Other Electronic Component Manufacturing, and Electronic Computer Manufacturing.

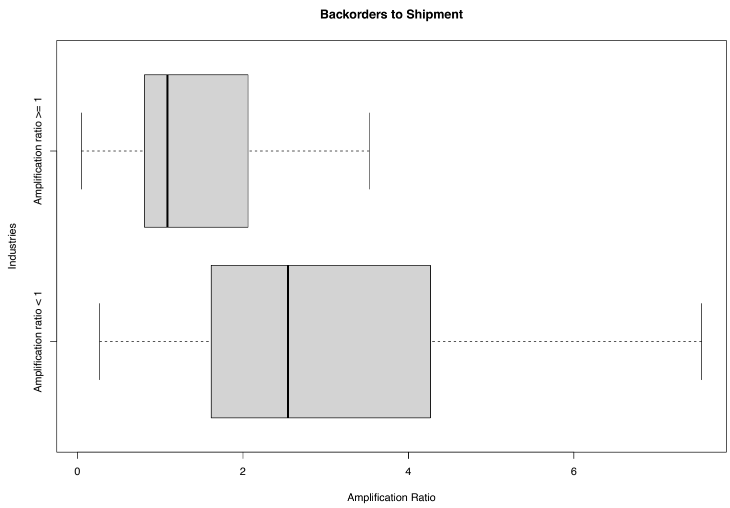

Figure 5.4: Box-plot: Amplification Ratio vs. Backorder Ratio

For these industries, backorders are simply queues that customers are willing to wait for customization. Electric Lighting Equipment Manufacturing, Ventilation, Heating, Air-Conditioning, and Refrigeration Equipment Manufacturing are among the other five industries with a positive correlation. These industries, expected to rely more on make-to-stock strategies, support the argument that BWE causes backorders.

The findings contradict Proposition 1, as many industries exhibit a negative correlation between inventory days and amplification, suggesting production smoothing rather than increased variance. Similarly, backorders show no consistent positive relationship with amplification, particularly in make-to-order industries where they function as queues rather than lost sales. Prior research also supports that BWE and production smoothing can coexist, further undermining the proposition. Thus, Proposition 1 is rejected.

5.2 Proposition 2: Overtime and Gross Job Loss positively correlated with the Amplification Ratio

The analysis shows that while overtime did not show a clear positive correlation with the amplification ratio, gross job loss, particularly in labor-intensive industries, suggested a positive correlation.

Figure 5.5 shows that 11 out of 29 industries indicate a negative correlation between amplification ratio and overtime, whereas there is only one industry positively correlated.

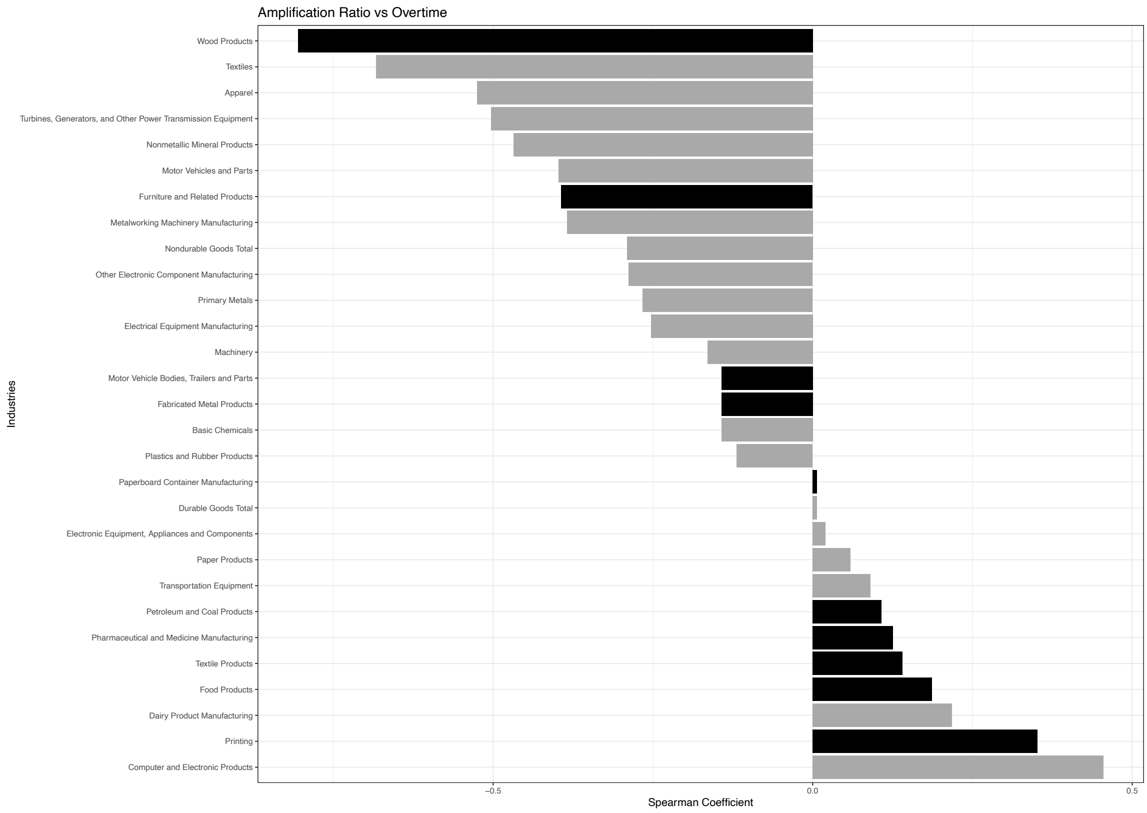

Figure 5.5: Bar-chart: Amplification Ratio vs. Overtime

Figure 5.6 Box- shows that industries that exhibit amplification tend to have less overtime. These results are interesting because it conflicts with Lee et al. (1997), which are not discussed in the literature. However, among negatively correlated industries, only two amplify demand variance. On the other hand, if significance levels are ignored, 6 out of 12 positively correlated industries bullwhips.

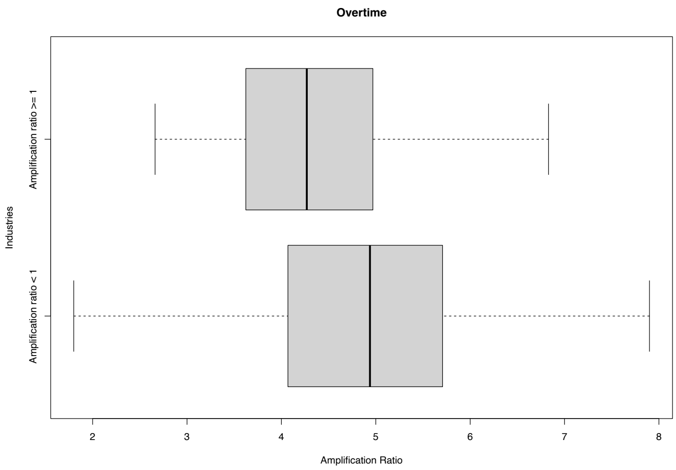

Figure 5.6: Box-plot: Amplification Ratio vs. Overtime

Nevertheless, Gross Job Loss results do not deny the positive association between Gross Job Loss and BWE, especially in more labor-intensive industries (e.g., Furniture and Related Products, Machinery, Wood Products, Printing). Figure 5.7 shows that among 18 industries, in 8, amplification ratio and gross job losses are positively correlated, and 5 of them bullwhip.

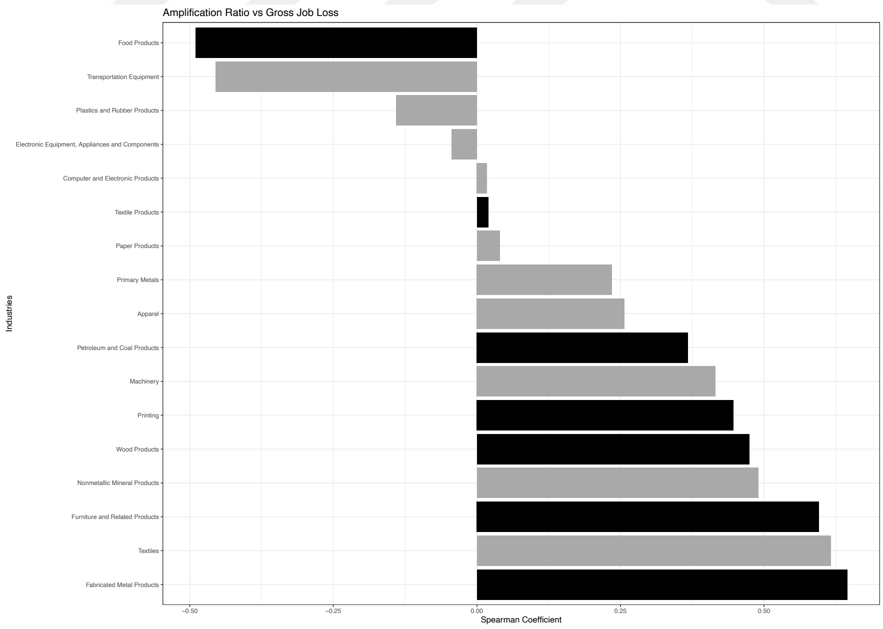

Figure 5.7: Bar-chart: Amplification Ratio vs. Gross Job Loss

Seven out of these eight industries are labor-intensive. Only two industries have a negative correlation, and none of them are labor intensive.

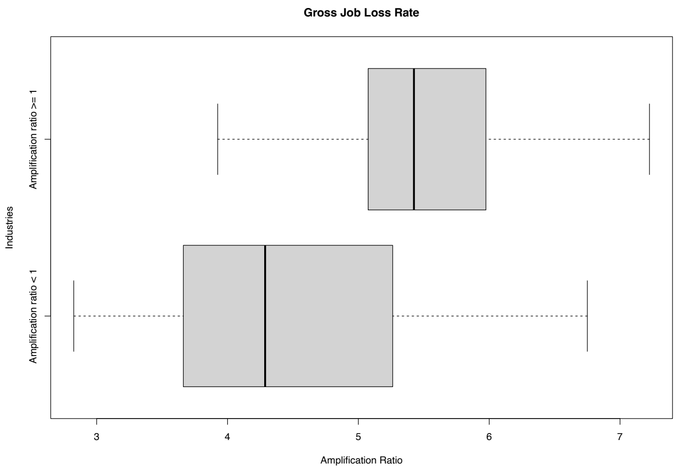

Figure 5.8: Box-plot: Amplification Ratio vs. Gross Job Loss

Figure 5.8 shows that bullwhipping industries also have higher gross job loss rates. So, results suggest that BWE may increase Gross Job Loss in labor-intensive industries and consequently negatively affect workers. This result concurs with Haltiwanger and Maccini (1988).

The findings partially reject the Proposition 2. The relationship between the amplification ratio and overtime, as most industries exhibit a negative correlation. Additionally, among negatively correlated industries, only a few amplify demand variance, further weakening the connection. Thus, the proposed positive correlation between BWE and overtime lacks sufficient support.

Conversely, the results partially accept the relationship between amplification and gross job loss, particularly in labor-intensive industries. The observed patterns support the argument that variance amplification contributes to workforce instability, especially in labor-dependent sectors.

5.3 Proposition 3: Amplification Ratio vs. Capacity Utilization

The correlation analysis of capacity utilization rates, considered both an underlying factor and a possible outcome of BWE, revealed that more industries tend to attenuate variance when capacity is lower. Figure 5.9 shows that out of 20 capacity utilization rate series, ten are negatively correlated with amplification ratio; among these 10, 5 industries bullwhip. The correlation coefficient is positive for the two industries.

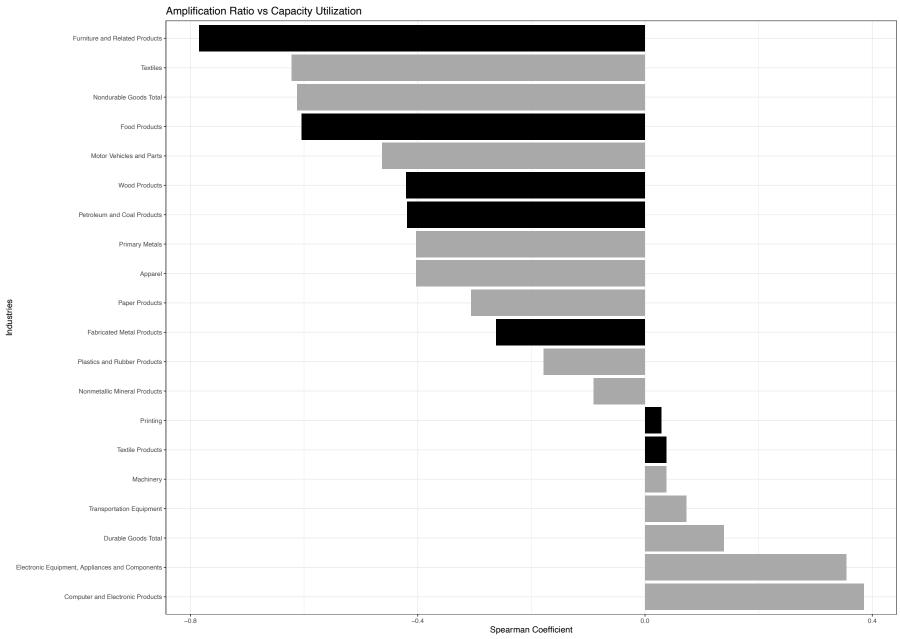

Figure 5.9: Bar-chart: Amplification Ratio vs. Capacity Utilization Rate

Figure 5.10 shows that the average capacity utilization rate is nearly equal for both amplifying and attenuating industries.

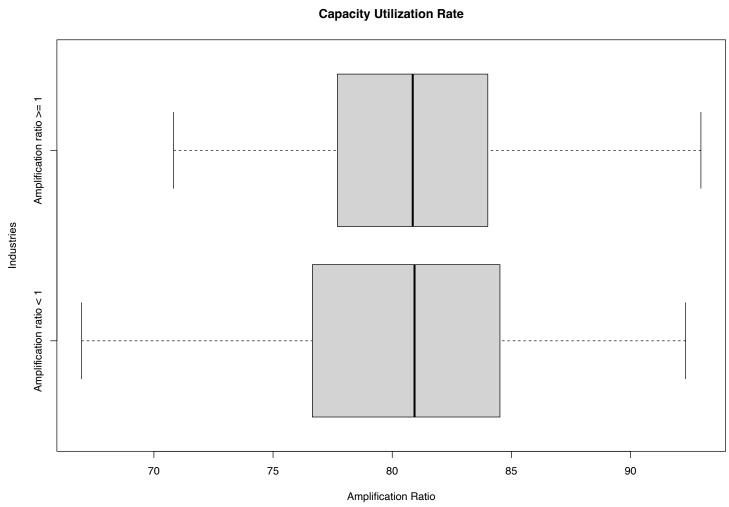

Figure 5.10: Box-plot: Amplification Ratio vs. Capacity Utilization Rate

There can be two explanations for these results. First, as Chen and Lee (2012) pointed out, limited capacity has a limiting effect on BWE. Since expensive investments are required to increase capacity, firms are needed to dampen the amplification. Positively correlated industries such as *Computer and Electronics and Electronic Equipment* and *Transportation Equipment* industries challenge this explanation, considering the need for expensive investments to adjust capacity. However, the ranks of these two industries in terms of average capacity utilization rate among 20 industries are 19 and 20, respectively. This relative slackness in utilization rates may give enough space for amplification.

Second, as Buchmeister *et al.* (2014) stated, better inventory policies yield better utilization rates. Both explanations confirm the results and conflict with Proposition 3; higher capacity utilization is a precursor for less bullwhip regardless if it is planned or imposed.

5.4 Complementary Analyses Between Variables

Although it is aimed to analyze industry-based figures, the pairwise correlations between variables are also analyzed to check an interrelation between variables. Table 5.1 presents the Pearson coefficient of correlation for each pair.

Table 5.1: Pearson Coefficient of Correlation Matrix

	Amplification Ratio	**Inventory Days**	**Overtime**	**Backorders to Shipments**	**Gross Job Loss**	**Capacity Utilization Rate**
Inventory Days	-0,19					
Overtime	-0,19	-0,12				
Backorders to Shipments	-0,32	-0,11	0,49			
Gross Job Loss	0,43	0,12	-0,62	-0,29		
Capacity Utilization Rate	-0,04	-0,04	0,27	-0,29	-0,54	
Inventory Days (log-differenced)	0,11	0,11	-0,07	-0,12	0,29	-0,02

All variables negatively affect the amplification ratio except for Gross Job Losses. It indicates a strong positive relationship with the amplification ratio, which concurs with Theoretical Proposition 2. Capacity utilization has the weakest correlation considering all time series. However, all correlations need to be addressed after disaggregating by industries to understand if industries behave the same against the amplification ratio.

A mild negative correlation is found between overtime and inventory days. This versus may seem irrelevant in the BWE context. However, it provides an insight into production smoothing. Production smooths if the firm uses its inventory to dampen demand shocks, increase its capacity to catch up with the demand, or take both measures. Results do not contradict the theory but may contribute to the discussion.

Again, if overtime is replaced with capacity utilization, findings support Buchmeister *et al.* (2014) study: an efficient production may increase inventory management performance.

Gross Job Loss has a strong negative association with overtime. This suggests that companies in industries in the U.S. are prone to do overtime instead of a hire-fire policy.

Interestingly, a mild negative association is found (Pearson coefficient, R = -0.46) between the seasonality ratio reported by Cachon *et al.* (2007) and the average inventory days of all manufacturing industries. This implies that companies have higher inventory turnover or shorter inventory days if seasonality is relatively more substantial. Considering seasonality as an exogenous variable, reverse causality cannot be made. Since Cachon *et al.* (2007) already showed that industries with more seasonality suffer BWE less, seasonality can be considered a regulating factor for inventory decisions.

5.5 Proposition 4: Reliance on imported inputs positively correlates with the Amplification Ratio

Pearson and Spearman correlation coefficients are calculated to explore pairwise associations between each industry's imported intermediate ratio and amplification ratio.

Amplification ratios are recalculated identically with Cachon *et* al. (2007). All amplification ratios except for four overlapping industries (Aluminum and Nonferrous Metal Products, Computer and Electronic Products, Motor Vehicle Bodies, Trailers and Parts, Ships and Boats and Turbines, Generators, and Other Power Transmission Equipment) are included in this study as they represent a more specific portion of a group of industries.

Both correlation methods are carried out for all pairs of variables in Table 5.2. Nevertheless, for most pairs, both methods gave similar results. Spearman and Pearson correlation coefficients are negative, -0.40 and -0.47, respectively (significant at 0.01).

Table 5.2: Amplification Ratio and Imported Intermediate Inputs

Industry	**Amplification Ratio**	**Imported Intermediate Ratio**
Aluminum and Nonferrous Metal Products	0,99	29%
Apparel manufacturing	0,57	18%
Basic Chemicals	0,74	13%
Beverage Manufacturing	**3,04**	15%
Communications Equipment, Total	0,35	33%
Computer and Electronic Products	0,14	24%
Computers and Related Products	0,28	30%
Construction Machinery Manufacturing	0,73	22%
Dairy Product Manufacturing	0,85	3%
Electric Lighting Equipment Manufacturing	0,43	16%
Electrical Equipment Manufacturing	0,70	21%
Electromedical, Measuring, and Control Instrument Manufacturing	0,48	19%
Electronic Computer Manufacturing	0,39	31%
Electronic Equipment, Appliances and Components	0,63	23%
Fabricated Metal Products	**1,01**	13%
Ferrous Metal Foundries	**1,21**	17%
Food Products	**1,32**	7%
Furniture and Related Products	**1,10**	18%
Grain and Oilseed Milling	**2,90**	5%
Household Appliance Manufacturing	0,69	16%
Industrial Machinery Manufacturing	0,23	14%
Iron and Steel Mills and Ferroalloy and Steel Products Manufacturing	0,81	18%
Leather and Allied Products	0,79	22%
Machinery	0,57	18%
Material Handling Equipment Manufacturing	0,33	22%
Meat, Poultry and Seafood Product Processing	**1,08**	7%
Metalworking Machinery Manufacturing	0,79	13%
Mining, Oil and Gas Field Machinery Manufacturing	**2,10**	17%
Motor Vehicle Bodies, Trailers and Parts	**1,04**	21%
Motor Vehicles and Parts	0,97	16%
Paper Products	0,99	16%
Paperboard Container Manufacturing	**1,40**	14%
Pharmaceutical and Medicine Manufacturing	**2,86**	6%
Printing	**1,59**	11%
Pulp, Paper, and Paperboard Mills	**1,20**	17%
Textile Products	**1,12**	13%
Textiles	0,57	13%
Tobacco Manufacturing	**3,09**	11%
Transportation Equipment	0,55	21%

Turbines, Generators, and Other Power Transmission Equipment	0,27	18%
HVAC, and Refrigeration Equipment Manufacturing	0,79	18%
Audio and Video Equipment Manufacturing	0,86	17%
Automobile Manufacturing	0,90	25%
Battery Manufacturing	**1,06**	33%
Computer Storage Device Manufacturing	0,20	23%
Farm Machinery and Equipment Manufacturing	0,88	22%
Heavy Duty Truck Manufacturing	**1,13**	24%
Light Truck and Utility Vehicle Manufacturing	0,97	25%

Figure 5.11 shows how industries are spread over amplification ratio and imported intermediate input. A linear regression line (OLS) and polynomial regression curve (LOESS) are also included to present the relationship between variables.

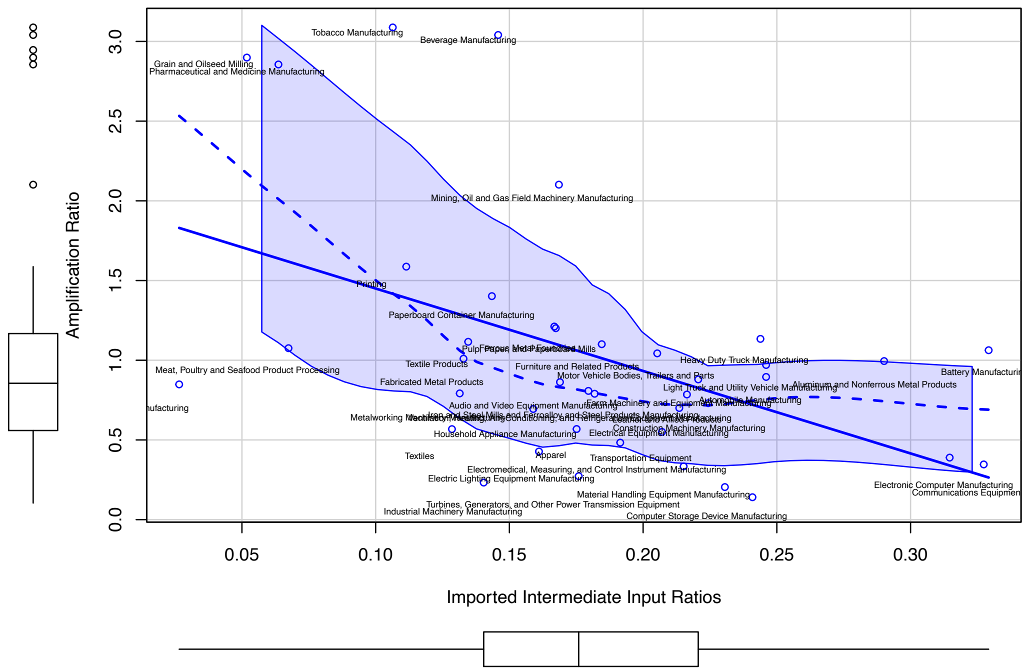

Figure 5.11: Amplification Ratio vs. Imported Intermediates Ratio

Industries with an imported intermediate ratio of over 20% exhibit reduced BWE. The amplification ratio for these industries, such as Heavy Duty Manufacturing and Battery Manufacturing, is slightly higher than one.

For further analysis, industries are grouped according to their classification in terms of the durability of their commodities. Figure 5.12 shows that non-durable industries are prone to have fewer imported intermediate inputs compared to durable industries. The only non-durable industry that exceeds 20% level is Leather and Allied Products industry. The time-sensitive demand nature of these industries can explain this barrier. For perishable commodities of food industries, firms may be unable to use imported inputs since the intermediates are not durable enough to withstand longer lead times. Djankov *et al.* (2010) indicate that delays significantly impact the import of time-sensitive goods, particularly perishable agricultural products, in terms of trade volume. Evans and Harrigan (2005) showed that textile and apparel industries are sensitive to distance, and imports are shifted to near-shoring these supply chains in the U.S. It is intuitive that time-sensitive products are prone to be supplied locally.

However, NAICS definition of non-durable goods covers all products having up to three years of shelf life. Printing and textile industries, which use commodities that do not expire fast, also fall behind the 20% level.

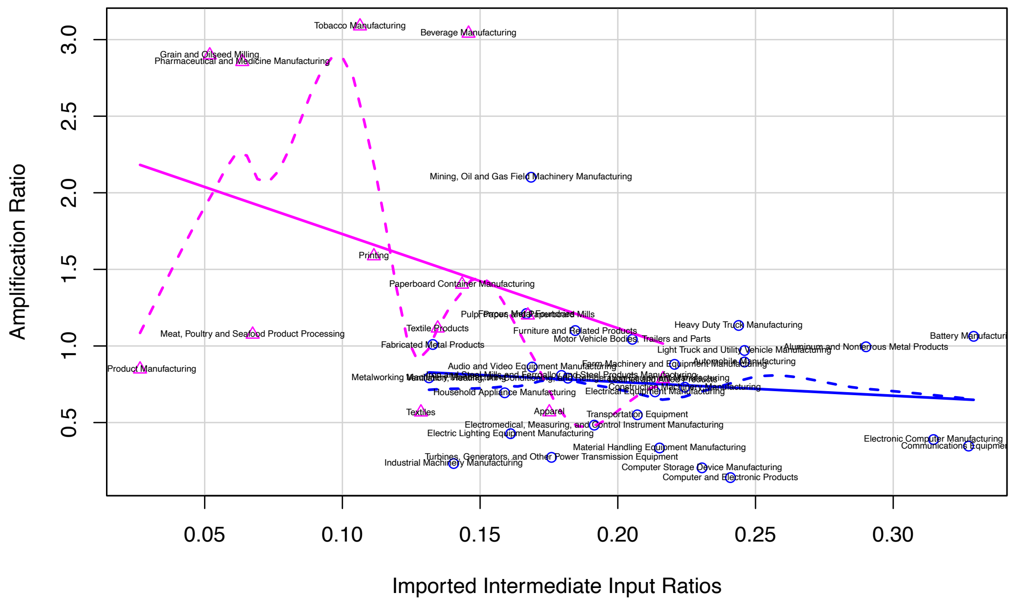

Figure 5.12: Amplification Ratio vs. Imported Intermediates Ratio Durable and Non-durable

Moreover, despite drivers that favor the local supply, BWE is more evident in less import-reliant industries. The amplification ratio within most durable industries exhibits values below one and shows diminished sensitivity to the influence of imported intermediate inputs. This raises the question of whether attenuating drivers prevail over amplifying drivers.

Table 5.3: Value per ton for matching industries

Industry	Amplification Ratio	Imported Intermediate Ratio	Thousand USD per Ton
Furniture and Related Products	1,1	0,18	5,43
Automobile Manufacturing	0,9	0,25	7,54
Machinery	0,57	0,18	9,21
Electronic Equipment, Appliances and Components	0,63	0,23	21,07
Transportation Equipment	0,55	0,21	35,97

To address this question, we propose utilizing the product's value relative to its weight as an indicator, which can explain the tendency of stakeholders toward information sharing and collaborative initiatives. Leveraging data sourced from the Freight Analysis Framework provided by the Bureau of Transportation Statistics, we aligned our statistical analyses with the assessed value of commodities transported within the U.S. Due to the restricted number of observations, a formal correlation test was not conducted. Nevertheless, Table 5.3 delineates an inverse relationship between the amplification ratio and the value per ton. This observed phenomenon is not mirrored in the case of non-durable products. Notably, pharmaceuticals, distinguished by their elevated value per ton, exemplify this trend, displaying one of the highest amplification ratios within this category.

Contrary to prior studies, the empirical analysis demonstrates that reliance on imported inputs does not exacerbate the BWE. Industries with higher imported intermediate input ratios tend to exhibit lower amplification, suggesting that longer transit times and import dependency do not necessarily contribute to demand variance amplification. Instead, factors such as product durability and local supply chain dynamics appear to play a more significant role in shaping amplification behavior. Thus, the proposition is rejected.

5.6 Proposition 5: The Trend in BWE Is Decreasing

The Mann-Kendall test, revealed that 30 of 77 industries had significant slopes, summarized in Table 5.4.

Table 5.4: Average amplification ratios for grouped years

Level	Industry	Feb 1992-1995	1996-2000	2001-2004	2005-2008	2009-2012	2013-2016	2017-2019	2020-2022 Aug	Statistic (Annual)	p.value (Annual)
M	Apparel	0,51	0,49	0,72	0,9	1,07	3,05	4,24	6,82	0,6731	0,000
M	Audio and Video Equipment	0,95	0,97	0,77	1,02	1,18	3,1	3,2	2,51	0,2688	0,035
M	*Battery Manufacturing*	*0,9*	*0,64*	*0,96*	*2,74*	*0,96*	*1,05*	*1,72*	*3,56*	*0,2817*	*0,027*
M	**Communications Equipment Manufacturing, Nondefense**	0,7	0,48	1,51	0,62	1	0,46	0,2	0,32	**-0,3247**	0,011
M	Computer Storage Device Manufacturing	0,43	0,15	0,13	0,12	0,34	0,36	0,7	0,39	0,3032	0,017
M	Electric Lighting Equipment Manufacturing	0,89	0,47	0,62	1,48	3,5	2,76	3,36	1,32	0,4882	0,000
M	**Electromedical, Measuring, and Control Instrument Manufacturing**	0,45	0,49	0,61	0,47	0,3	0,11	0,24	0,4	**-0,3032**	0,017
M	Electronic Computer Manufacturing	0,31	0,48	0,66	1,37	1,6	0,65	1,69	1,87	0,4495	0,000
M	Furniture and Related Products	1,35	1,3	2,96	1,96	2,05	1,91	3,06	1,92	0,3118	0,014
M	**Household Appliance Manufacturing**	2,08	1,28	0,9	3,11	2,9	1,11	1,05	0,46	**-0,2559**	0,045
M	*Industrial Machinery Manufacturing*	*1,4*	*1,01*	*0,84*	*0,75*	*1,38*	*0,88*	*2,04*	*4,21*	*0,2559*	*0,045*
M	Iron and Steel Mills and Ferroalloy and Steel Product Manufacturing	2,18	0,63	1,07	0,79	1,6	1,33	1,27	3,1	0,2946	0,021
M	Leather and Allied Products	0,58	0,47	0,62	1,13	1,39	1,67	1,15	1,13	0,3204	0,012
M	***Light Truck and Utility Vehicle Manufacturing***	*1,12*	*0,92*	*0,93*	*1,02*	*0,87*	*0,75*	*0,65*	*1,39*	***-0,5054***	*0,000*
M	**Miscellaneous Products**	0,68	0,82	0,49	0,36	0,47	0,4	0,35	0,22	**-0,3419**	0,007
M	Other Electronic Component Manufacturing	1,59	0,92	1,5	0,52	0,88	3,12	6,57	7,04	0,3892	0,002
M	Pharmaceutical and Medicine Manufacturing	1,27	2,32	3,67	5,79	4,63	3,76	8,4	3,75	0,3806	0,003
M	**Search and Navigation Equipment, Defense**	3,77	0,71	1,68	0,53	0,64	3,99	0,62	0,6	**-0,2516**	0,049
M	Textile Mills	0,51	0,53	0,68	0,99	1,25	1,07	0,89	0,76	0,4065	0,001
M	Textile Products	1,26	0,95	1,04	1,37	1,73	2,04	1,77	1,64	0,3677	0,004
M	**Tobacco Manufacturing**	2,75	8,03	2,02	4,13	1,34	2,8	2,03	1,72	**-0,3075**	0,016
M	**Transportation Equipment**	0,89	0,85	0,74	0,65	0,63	0,7	0,52	0,82	**-0,4323**	0,001
R	**Building Mat. and Garden Equip. and Supplies Dealers**	0,97	0,96	0,9	1,1	0,77	0,66	0,87	0,72	**-0,3247**	0,011
R	***Motor Vehicle and Parts Dealers***	*2,58*	*2,49*	*0,96*	*1,61*	*1,29*	*1,33*	*0,87*	*2,26*	***-0,2817***	*0,027*
W	**Chemicals and Allied Products**	1,45	1,5	1,22	1,29	1,19	0,82	0,82	1,08	**-0,2645**	0,038
W	**Drugs and Druggists' Sundries**	5,42	3,59	3,39	2,49	3,23	2,75	2,78	2,06	**-0,4022**	0,002
W	**Machinery, Equipment, and Supplies**	1,1	1,59	1,05	1,21	0,89	0,64	0,66	0,7	**-0,4495**	0,000
W	**Metals and Minerals, Except Petroleum**	2,41	1,17	0,89	0,99	1,46	0,89	0,9	0,95	**-0,3161**	0,013
W	**Miscellaneous Non-durable Goods**	1,71	1,25	1,38	1,07	0,8	0,62	0,85	1,35	**-0,3204**	0,012
W	*Petroleum and Petroleum Products*	*1,15*	*1,49*	*1,34*	*1,52*	*1,79*	*1,5*	*2,28*	*4,07*	*0,3677*	*0,004*

The statistics column in Table 5.4 reflects the direction of the trend, which was negative for all wholesale and most retail industries except petroleum and petroleum products. In contrast, 14 manufacturing industries had a positive trend, while eight were negative. Overall, half of the significant trends were positive, indicating that not all industries exhibit a negative trend in the amplification ratio.

The results do not support a consistent decreasing trend in the BWE across industries. While wholesale and most retail industries, except petroleum and petroleum products, exhibit a negative trend, the findings for manufacturing industries are mixed. A significant portion of manufacturing industries show an increasing trend, and nearly half of all significant trends are positive. These results indicate that the amplification ratio is not uniformly declining across sectors. Thus, the proposition is rejected.

5.7 Proposition 6: The COVID-19 pandemic has significantly increased BWE within global supply chains

5.7.1 Forecast Method Performance Comparison

Table 5.5 shows the comparative analysis of forecasting models across two distinct periods, shown in 2016-2019 (pre-COVID-19) and 2020-2023(during and post-COVID-19), respectively, with the average MAPE and MAPE variances of 77 industries both for Demand and Inventory series.

Table 5.5: Forecasting Results

		2016-2019		2020-2023	
	Method	MAPE (%)	MAPE_Var (%)	MAPE (%)	MAPE_Var (%)
Demand	SARIMA	9.05	0.74	21.46	3.78
	Prophet	12.63	1.80	31.65	13.75
	RNN	4.17	0.23	6.64	0.73
	LSTM	**4.07**	**0.22**	**6.63**	**0.72**
Inventory	SARIMA	9.58	1.20	23.42	7.74
	Prophet	12.64	1.68	31.78	10.86
	RNN	1.86	0.02	2.60	0.03
	LSTM	**1.83**	**0.02**	**2.57**	**0.03**

For Demand Forecasting, SARIMA and Prophet show significant increases in MAPE (21.46 % and 31.65%) during the COVID-19 period, indicating a decline in forecast accuracy under the unpredictable conditions brought about by the pandemic. The increase in MAPE_Var (3.78% and 13.75%) for these methods suggests that their forecasts became less consistent over time. In contrast, RNN and LSTM both demonstrate superior performance with much lower increases in MAPE from the pre-COVID-19 (4.17% and

4.07%) to the COVID-19 period (6.64% and 6.63%) compared to SARIMA and Prophet. They also maintain lower MAPE_Var values (0.73% and 0.72%), suggesting more stable forecasts despite the changing conditions.

Similar to demand forecasting, SARIMA, and Prophet exhibit larger errors and variances during COVID-19 (23.42% and 31.78%), reflecting challenges in adapting to the volatile environment. RNN and LSTM again outperform SARIMA and Prophet, maintaining remarkably low MAPE (2.60% and 2.57%) and MAPE_Var values even during the pandemic. This indicates high robustness and reliability in more complex and nonlinear scenarios.

Overall, RNN and LSTM show outstanding resilience to the disruptions caused by COVID-19, maintaining high accuracy and consistency in their forecasts for demand and inventory. Traditional models like SARIMA and Prophet struggled more with the volatility and irregularities brought about by the pandemic, leading to higher errors and inconsistencies. The data suggests that for handling more complex, nonlinear dynamics (especially in unpredictable scenarios like a pandemic), models like RNN and LSTM are preferable due to their ability to adapt and maintain performance.

Figure 5.13 and Figure 5.14 represent the MAPE distribution of all the forecasting models used in 77 industries for the Demand and Inventory Series, respectively.

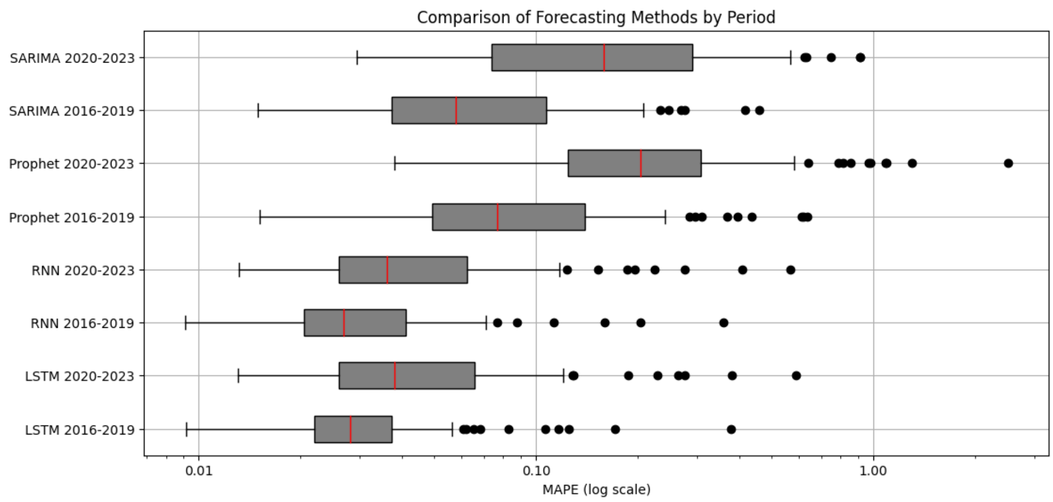

Figure 5.13: MAPE Boxplots (Demand)

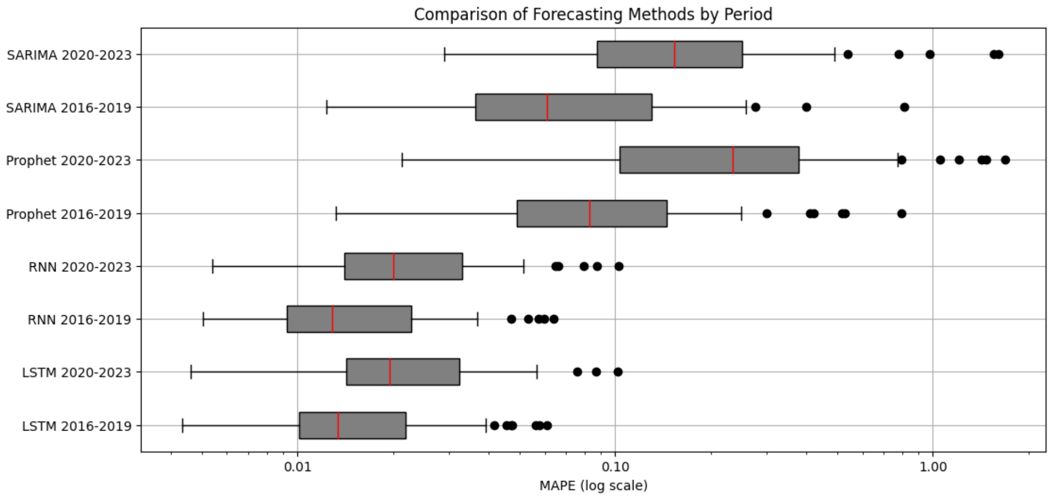

Figure 5.14: MAPE Boxplots (Inventory)

The performance of SARIMA significantly deteriorated in the 2020-2023 period, with a substantial increase in MAPE. RNN and LSTM showed consistent performance, although the error metrics slightly increased in the later period. This minimal increase suggests that these models maintained their predictive accuracy relatively well. The small rise in MAPE might be attributed to changes in underlying data patterns or external economic factors not captured in the model training phase. Prophet exhibited the most significant increase in error, with drastically higher MAPE values in the 2020-2023 period.

In conclusion, while RNN and LSTM stand out as superior forecasting tools for demand and inventory series, the LSTM forecasts edge slightly over RNN regarding consistency and reliability. Therefore, LSTM is selected as the baseline predictor for the Bullwhip Analysis stage.

5.7.2 Amplification Ratio Comparison

Figure 5.15 compares the actual amplification ratios against the predicted ones for 2020-2023, calculated with Equation 2. Whereas Actual Amplification Ratio is based on test data, LSTM Amplification Ratio is based on forecasted demand and inventory series. The scatter diagram highlights the manufacturer, wholesaler, and retailer supply chain stages in different colors. In addition, each point in the scatter diagram is labeled with the industrial code. The diagram is divided into four zones, according to the calculated amplification ratios for the actual and predicted series, where values below 1 indicate no evidence for BWE (-), and values greater than 1 indicate BWE (+).

In the "Accurately Forecasted Bullwhip" zone (+,+), both forecasted and actual series indicate a BWE. "False Positive Bullwhip" (+,-) is populated with forecasts predicting a BWE, but the actual values did not substantiate this. Given that BWE is a distinct supply chain phenomenon, it should not be artificially generated by predictors. Nonetheless, forecasting errors can lead to inaccurately amplified variance. To ensure the robustness of a predictor, this category should ideally contain the fewest instances, as is observed in our study.

In the "Accurately Forecasted No Bullwhip" zone (-,-), both forecasted and actual values concur on the absence of a BWE, indicating that these industries are less frequently subjected to BWE systematically. "False Negative Bullwhip" (-,+), on the other hand, is comprised of forecasts that failed to predict a BWE that was indeed evident in the actual values. Industries within this category are a focal point of our analysis, particularly those potentially more impacted by COVID-19.

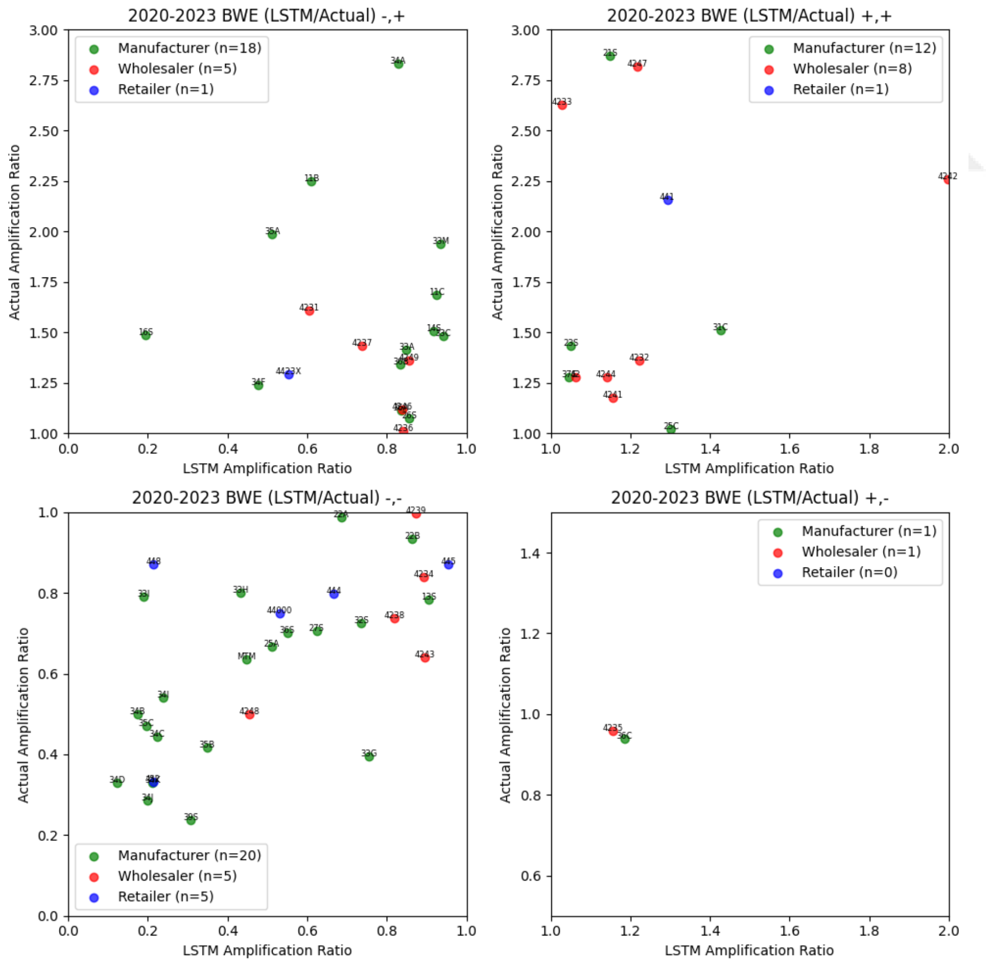

Figure 5.15: Amplification Ratios (Forecasted vs Actual)

Table 5.6: summarizes the Figure 5.15 with aggregate results. The percentages for each stage are given next to the number of instances.

Table 5.6: LSTM Forecast and Actual BWE and Stages

Stage	Accurately Forecasted Bullwhip (+,+)	Accurately Forecasted No Bullwhip (-,-)	False Positive Bullwhip (+,-)	False Negative Bullwhip (-,+)
Manufacturer	12 (23.5%)	20 (39.2%)	1 (0.02%)	18 (35.3%)
Retailer	1 (14.2%)	5 (71.4%)	-	1 (14.2%)
Wholesaler	8 (42.1%)	5 (26.3%)	1 (5.3%)	5 (26.3%)
Total	21 (27.3%)	30 (39.0%)	2 (2.6%)	24 (31.2%)

Manufacturers display a higher frequency of both Accurately Forecasted Bullwhip (12; 23.5%) False Negatives (-,+) (18; 35.3%) incidents, 58.8% of all manufacturing industries, suggesting a greater susceptibility to both the endogenous and external BWE. This aligns with the traditional understanding of the BWE, where demand variance amplification tends to propagate upstream, impacting manufacturers the most due to their position in the supply chain.

Wholesalers appear more prone to the BWE than retailers, as evidenced by their higher rates of False-negative incidents (5; 26.3%). This indicates that wholesalers are not immune to exogenous shocks, reflecting their intermediate position in the supply chain subjecting them to upstream supply variations and downstream demand fluctuations. However, showing moderate levels of both correct and missed predictions. Their intermediary position may offer better insights into supply and demand trends, helping mitigate some BWEs.

Retailers appear the least affected by the BWE with fewer instances (5; 71.4%) of the BWE, with a majority exhibiting variance smoothing. This observation is consistent with the BWE's theoretical framework, as retailers, being closer to the consumer end of the supply chain, typically experience less demand amplification.

The aggregated data reveal a general trend of higher bullwhip susceptibility upstream with manufacturers and a moderate impact on wholesalers. The lower impact on retailers highlights the effectiveness of proximity to the final consumer in mitigating the BWE. These insights emphasize the need for targeted strategies to address and mitigate the BWE, particularly for upstream supply chain actors.

The findings strongly support the proposition that the COVID-19 pandemic has significantly increased the BWE within global supply chains. The comparative analysis of pre-pandemic and pandemic periods reveals substantial disruptions in demand patterns, inventory management, and forecasting accuracy across industries. Higher amplification ratios observed during COVID-19 indicate intensified demand variability, particularly in upstream supply chain stages, aligning with traditional BWE dynamics. The forecasting results further confirm that industries faced greater volatility, with traditional models struggling to adapt, while advanced neural network-based models like LSTM demonstrated more stability. Additionally, the study highlights the increased susceptibility of manufacturers and wholesalers to BWE, reinforcing that supply chain instability intensified under pandemic-induced external shocks. These empirical insights confirm that COVID-19 amplified the BWE, validating the proposition.

5.7.3 Susceptibility of Industries Against Bullwhip Effect in COVID-19

The industries where the LSTM model did not predict the BWE (-,+) where it exists suggest a high susceptibility to this phenomenon, particularly during unpredictable pandemic conditions. These industries experienced demand surges, demand stops, or supply disruptions due to labor or raw material shortages that can exacerbate the BWE, summarized in Table 5.7. Sample plots of demand series and corresponding predictions of forecast methods are given for each stage to illustrate the comparison between predicted and actual patterns.

Table 5.7: False Negative Bullwhip Industries

Code	Industry	Demand Shock	Supply Shock
11C	Meat, Poultry, and Seafood Product Processing (M)	Less affected ⟷	Negative ↓
26S	Plastics and Rubber Products (M)		
33A	Farm Machinery and Equipment Manufacturing (M)		
33E	Industrial Machinery Manufacturing (M)		
35A	Electric Lighting Equipment Manufacturing (M)		
35D	Battery Manufacturing (M)		
14S	Textile Products (M)	Negative ↓	Negative ↓
15S	Apparel (M)		
16S	Leather and Allied Products (M)		
33C	Construction Machinery Manufacturing (M)	Negative ↓	Less affected ⟷
11B	Dairy Product Manufacturing (M)	Positive ↑	Negative ↓

33M	Material Handling Equipment Manufacturing (M)		
34A	Electronic Computer Manufacturing (M)		
34F	Audio and Video Equipment (M)		
34H	Other Electronic Component Manufacturing (M)		
36A	Automobile Manufacturing (M)		
36B	Light Truck and Utility Vehicle Manufacturing (M)		
31A	Iron and Steel Mills and Ferroalloy and Steel Product (M)	Less affected ↔	Negative ↓
4246	Chemicals and Allied Products (W)	Less affected ↔	Negative ↓
4237	Hardware, and Plumbing and Heating Equipment and Supplies (W)		
4231	Motor Vehicle and Motor Vehicle Parts and Supplies (W)	Negative ↓	Negative ↓
4236	Household Appliances and Electrical and Electronic Goods (W)	Positive ↑	Negative ↓
4249	Miscellaneous Nondurable Goods (W)		
4423X	Furniture, Home Furn., Electronics, and Appliance Stores (R)	Negative ↓	Negative ↓

Eighteen manufacturer industries where the model did not predict a BWE that actually occurred are discussed below with the support of related literature. Demand for Meat, Poultry, and Seafood, Plastics and Rubber Products industries remained stable, driven by consistent consumer needs. However, the supply was challenged by outbreaks of COVID-19 in processing plants, leading to stringent health and safety measures to protect workers and maintain production levels Brinca *et al.* (2021); Ramsey *et al.* (2021). Industrial Machinery, Farm Machinery and Equipment, Electric Lighting Equipment, and Battery Manufacturing industries were relatively unaffected as many sectors reliant on these industries and demand for agriculture continued during lockdowns, whereas supply suffered labor shortages Maria del Rio-Chanona *et al.* (2020).

In the Textile Products, Apparel, and Leather and Allied Products Manufacturing industries, a sharp decline in demand was observed in fashion and luxury goods saw reduced consumer spending. On the supply side, production facilities faced operational restrictions due to disruptions caused by factory closures and restrictions in manufacturing hubs Cho and Saki (2022); Seidu *et al.* (2023).

Construction Machinery Manufacturing industry experienced a decline in demand as non-essential construction projects were suspended in response to global lockdowns and social distancing measures (Brinca et al., 2021). These restrictions severely impacted the industry's ability to continue operations, leading to a drop in demand for construction machinery. Despite these challenges, the supply side was relatively more stable but not without disruptions, primarily due to reduced labor availability as workers were affected by health concerns and restrictions on movement Ghansah and Lu (2023).

As Liu and Rabinowitz (2021) study details, the imposition of stay-at-home orders and the closure of key institutional buyers like schools and restaurants shifted consumption patterns from dining out to home consumption, substantially increasing demand for Dairy Products. Demand for Material Handling Equipment and Light Truck and Utility Vehicle Manufacturing industries rose due to the expansion of e-commerce and increased automation needs in logistics operations Herold *et al.* (2021). Audio and Video Equipment Manufacturing, Electronic Computer Manufacturing, and Other Electronic Component industries saw a surge in demand as consumers sought entertainment while confined to their homes, the surge in remote working and online schooling and eventually increasing need for electronic components in consumer electronics Nikolopoulos *et al.* (2021). However, these industries faced supply disruptions due to the closure of manufacturing facilities, international logistics, and component shortages Handfield *et al.* (2020); Walmsley *et al.* (2023).

Automobile Manufacturing initially experienced a dip in demand due to economic uncertainties (Figure 5.16). Demand subsequently increased as consumers began to prefer private vehicles over public transportation to minimize exposure to the virus Basu and Ferreira (2021); Parker *et al.* (2021). The supply side in this sector was hampered by disruptions across the global supply chain and labor shortages Belhadi *et al.* (2021).

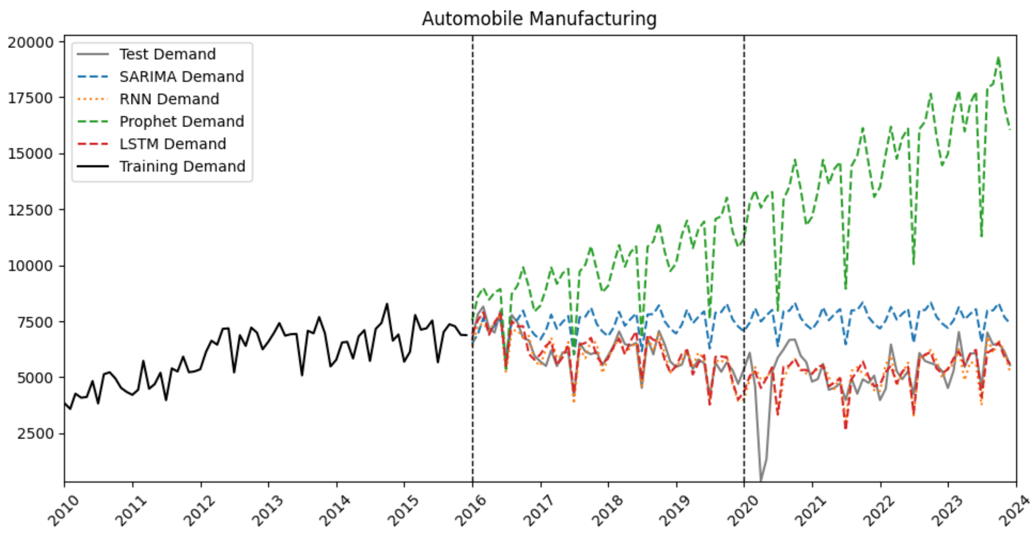

Figure 5.16: Demand Series for Automobile Manufacturing

While demand impacts were minimal, the Iron and Steel Mills and Ferroalloy and Steel Product Manufacturing industry struggled with supply chain disruptions from labor shortages, highlighting the need for more flexible production strategies Asadi *et al.* (2023).

Five wholesaler industries where the BWE was observed, despite the did not predict it. Chemicals and Allied Products Wholesaling industry, critical in supplying essential materials across various sectors, faced unprecedented challenges due to global lockdown measures and border closures that disrupted supply chains and raw material availability. Demand patterns were notably altered. There was a surge in demand for chemicals necessary in pharmaceuticals and healthcare, including disinfectants and sanitizers Abedsoltan (2023).

The Household Appliances and Electrical and Electronic Goods Wholesaling industry saw increased demand after an initial shock for home appliances due to more time spent at home, as shown in Figure 5.17. However, supply chains were challenged by manufacturing and international logistics disruptions, highlighting the need for supply chain visibility and flexibility Nikolopoulos *et al.* (2021).

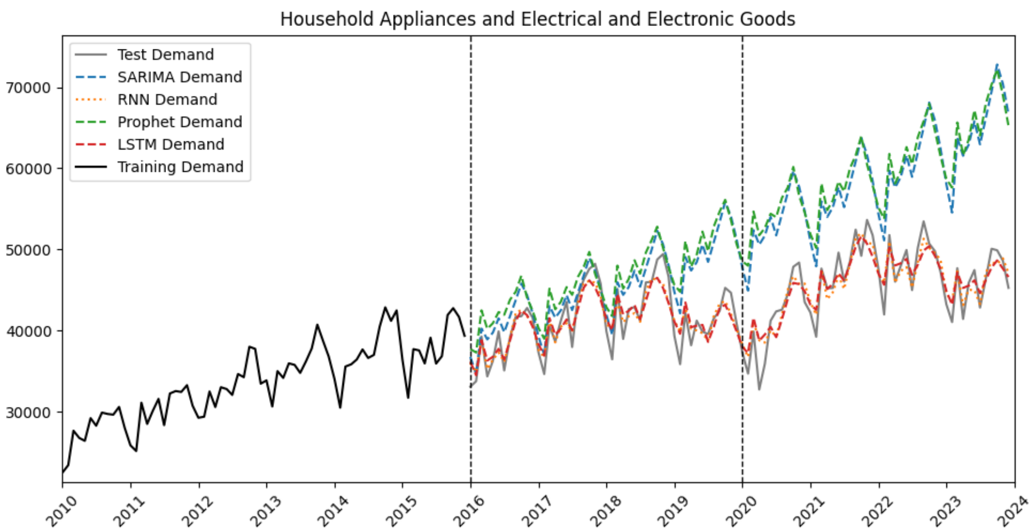

Figure 5.17: Demand Series for Household Appliances and Electrical and Electronic Goods Wholesaling

The Miscellaneous Nondurable Goods Wholesaling industry experienced demand spikes from panic buying, particularly for hygiene products, while supply was constrained by disruptions in manufacturing and logistics Paul and Chowdhury (2020).

Motor Vehicle and Motor Vehicle Parts and Supplies Wholesaling suffered a demand decrease due to reduced car usage due to lockdowns. At the same time, supply issues stemmed from disruptions in global automotive manufacturing Handfield *et al.* (2020).

Retailers appear the least affected by the BWE, which could be attributed to their end-of-chain position, allowing for more direct consumer demand assessments and rapid inventory and ordering strategy adjustments.

Furniture, Home Furnishings, Electronics, And Appliance Retailing industry experienced an initial demand shock in demand as consumers did not prioritize home improvement during lockdowns, as shown in Figure 5.18 Verhoef *et al.* (2023). At the same time, the supply faced challenges due to the closures of retail outlets and disruptions in global manufacturing and shipping.

During the pandemic, diverse industries experienced significant shifts in demand and supply dynamics, reflecting the global crisis's broad impact and the varied nature of industrial operations.

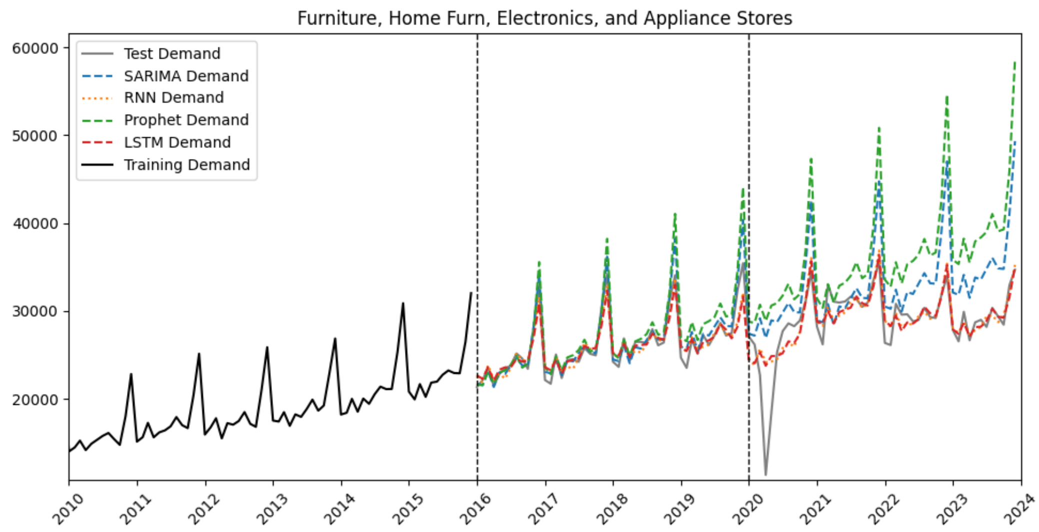

Figure 5.18: Demand Series for Furniture, Home Furnishings, Electronics, and Appliance Retailing

The results suggest that the COVID-19 pandemic reshaped consumer behavior and demand across various sectors and highlighted vulnerabilities in global supply chains to trigger BWE. Supply disruptions have varied in nature and intensity across different industries, making it essential to study the BWE within a pandemic context to understand better how such external shocks can influence supply chain dynamics. In this context, the unprecedented disruptions brought forth by the pandemic provided a novel context for its examination, particularly how external shocks can exacerbate it.

6. DISCUSSION AND MANAGERIAL IMPLICATIONS

The findings of this study challenge simplistic assumptions regarding the BWE in industry-level data, providing nuanced insights into the differential impact of contributing factors across industries. While previous literature has established a consensus on certain drivers of BWE, this research reveals that industries experience BWE effects differently, necessitating a more tailored approach to supply chain management and mitigation strategies.

6.1 Reevaluating the Adverse Effects of BWE: Industry-Specific Insights (Propositions 1, 2, and 3)

The analysis of industry-level data suggests that not all traditionally accepted drivers of BWE affect industries equally. While inventory strategies and production capacity appear to smooth demand fluctuations, BWE do not significantly contribute to backorder levels, contradicting conventional assumptions. The only factor that strongly correlates with BWE across industries is job loss, suggesting that labor volatility plays a central role in demand amplification. These findings have critical managerial implications:

- Benchmarking Best Practices: Industries that have successfully mitigated BWE through inventory smoothing and capacity management should serve as models for best practices. Companies in more volatile sectors should study and adapt the buffering and flexibility strategies of industries where these mechanisms have proven effective.
- Labor Stability as a Mitigation Strategy: Given the strong link between job loss and BWE, firms should prioritize workforce stability, particularly in industries prone to high demand variability. Adopting flexible employment models, cross-training, and workforce-sharing programs can help reduce labor-induced demand

fluctuations.

- Rethinking Traditional Presumptions: Managers should not assume that common BWE factors apply uniformly across all industries. Instead, firms should conduct industry-specific risk assessments to identify their primary BWE drivers and implement targeted solutions rather than generic supply chain strategies.

6.2 The Role of Importation and Supply Chain Collaboration (Proposition 4)

Contrary to prior studies, the empirical analysis conducted in this research demonstrates that importation and import-related transit times do not appear to exacerbate the BWE. The study reveals that the attenuating forces within the causal diagram seem to outweigh numerous amplifying factors. This intriguing finding implies that other variables and dynamics within the supply chain, such as information sharing and collaboration initiatives, may play a pivotal role in mitigating BWE.

Manufacturers relying more on imported intermediates may have less flexibility in propagating variance, necessitating more efficient planning for middle-term supply requirements to mitigate BWE. Moreover, foreign trade operations often necessitate a high level of trust and partnership between firms, leading to the establishment of long-term relationships compared to domestic transactions. These long-term relationships are a potential positive driver of collaboration, particularly regarding information sharing. Thus, the level of trust and the nature of the partnerships forged in international trade are linked to enhanced information-sharing practices, contributing to the overall resilience of the supply chain against BWE. For managers, this insight underscores the importance of prioritizing collaboration over near-shoring:

- Investing in Digital Collaboration Tools: Instead of focusing solely on reducing lead times through near-shoring or reshoring, firms should prioritize enhancing information transparency with suppliers. Cloud-based platforms, real-time data sharing, and blockchain-enabled supply chain visibility can mitigate BWE by improving demand forecasting and reducing uncertainty.
- Developing Stronger Supply Chain Partnerships: Firms should shift from transactional supplier relationships to long-term partnerships that emphasize trust, data exchange, and coordinated planning. Supplier collaboration programs, such as Vendor-Managed Inventory (VMI) and joint forecasting initiatives, can

significantly reduce demand amplification.

- Rethinking Lead Time as a Strategic Variable: Instead of viewing longer lead times as inherently negative, managers should recognize that industries with strong collaboration mechanisms have successfully managed long supply chains without experiencing significant BWE effects. This suggests that improving coordination can be a more effective solution than simply reducing lead times.

6.3 The Persistent Trend of BWE and the Role of Emerging Technologies (Proposition 5)

Despite significant advancements in information technology, the study finds no evidence that BWE prevalence is decreasing over time. This suggests that conventional digital solutions have not been effective in mitigating demand variability, necessitating new approaches. For industry leaders, this calls for:

- Increased Adoption of AI-Driven Supply Chain Solutions: Traditional Enterprise Resource Planning (ERP) systems and demand forecasting software may not be sufficient to combat BWE. Instead, companies should leverage advanced AI-driven predictive analytics, reinforcement learning algorithms, and dynamic pricing models to improve demand forecasting accuracy.
- Enhancing Decision-Making with AI-Powered Collaboration Tools: While traditional IT systems facilitate data storage and communication, they often fail to optimize real-time decision-making. Firms should implement AI-enhanced collaborative platforms that allow real-time adjustments to supply chain disruptions based on predictive analytics and external market signals.
- Developing Cross-Industry Knowledge Sharing: Given that some industries have naturally lower BWE susceptibility, firms in high-BWE sectors should actively engage in industry collaborations and cross-sector learning initiatives. Joint research consortia, data-sharing partnerships, and collaborative training programs can help firms learn from industries that have successfully reduced demand volatility.

6.4 The Impact of COVID-19 on BWE: Managing Susceptibility to Demand Shocks (Proposition 6)

The study highlights that some industries are inherently more susceptible to BWE than others, particularly during periods of global disruption such as the COVID-19 pandemic. The research suggests that firms operating in high-BWE industries should proactively identify their susceptibility and implement best practices. Key managerial recommendations include:

Industry-Specific Risk Management Frameworks: Industries with high BWE susceptibility should implement more resilient supply chain strategies that emphasize demand sensing, real-time inventory tracking, and dynamic risk modeling. Scenario-based planning and multi-tier supplier risk assessments can help mitigate disruptions.

Adopting Best Practices from Resilient Industries: Firms in vulnerable industries should benchmark supply chain resilience strategies from sectors that have demonstrated strong BWE resistance. This includes diversifying supplier networks, strengthening logistics partnerships, and leveraging automated decision-support systems.

Building More Adaptive Supply Chain Structures: Rigid supply chains tend to exacerbate BWE during demand shocks. Managers should shift toward more flexible supply chain designs, including agile manufacturing, distributed production networks, and real-time inventory optimization models.

6.5 A Strategic Shift in BWE Mitigation

This study reveals that conventional supply chain assumptions about BWE require reconsideration. Industry-level data suggests that not all commonly cited drivers of BWE affect industries equally, and collaboration, rather than near-shoring, is the most effective countermeasure. Moreover, despite advancements in information technology, BWE has not declined over time, indicating the need for AI-driven forecasting methods and real-time collaboration tools. For managers and policymakers, the key takeaways from this study are:

- Industry-Specific BWE Mitigation: Rather than applying generalized supply chain strategies, firms should adopt customized solutions based on sector-specific insights.

- Collaboration Over Location-Based Strategies: Near-shoring alone does not mitigate BWE. Instead, firms should focus on enhancing collaboration, information sharing, and supply chain transparency.
- Technology-Driven Risk Management: Traditional IT systems have not significantly reduced BWE. To achieve meaningful improvements, firms must integrate AI-based forecasting, predictive analytics, and adaptive supply chain strategies.
- Proactive Risk Mitigation in High-BWE Sectors: Industries that are more susceptible to demand shocks should take a proactive approach to risk management, leveraging best practices from resilient sectors.

By implementing these strategic insights, firms can move beyond reactive supply chain management and actively mitigate demand amplification, ultimately fostering more stable and efficient global supply chains.

7. ADVANCED FORECASTING METHODS TO MITIGATING THE BULLWHIP EFFECT

The findings from industry-level analysis underscore the critical need for advanced forecasting solutions to mitigate the persistent effects of the BWE. While conventional supply chain strategies—such as inventory buffering, capacity adjustments, and lead-time reduction—have demonstrated limited success in curbing demand amplification, the study highlights forecasting accuracy and real-time collaboration as the most effective levers for controlling order variability. However, despite significant advancements in information technology, BWE has not diminished over time, suggesting that existing forecasting methods remain inadequate.

Recognizing this gap, this study proposes a novel forecasting approach that integrates AI-driven predictive models, dynamic learning mechanisms, and multi-source data fusion to improve demand accuracy and enhance supply chain resilience. The development of these advanced forecasting methods is directly motivated by the need to counteract the limitations of traditional demand planning techniques, enabling firms to anticipate fluctuations more accurately and respond proactively to supply chain disruptions. The following section introduces state-of-the-art forecasting methodologies, demonstrating how machine learning, deep learning, and hybrid statistical techniques can provide robust, adaptable, and data-driven solutions to mitigate the adverse effects of BWE across industries.

The central objective of this study was to assess the effectiveness of one-step and multi-step (h-step-ahead) forecasting methods in mitigating the BWE. Unlike traditional methods such as Moving Average (MA) and Exponential Smoothing (ES), ML techniques leverage advanced computational capabilities to predict demand patterns more accurately under various conditions of autocorrelation, seasonality, and variability. This research aims to:

Evaluate the forecasting accuracy of LSTM and LightGBM against traditional methods under autoregressive demand processes and real-world datasets (M5 competition).

Examine the impact of forecasting methods on key supply chain metrics such as the amplification ratio, inventory variance, order fulfillment rate, and average end inventory.

Provide actionable insights into the potential of ML-based methods to better manage supply chain variability and reduce BWE.

7.1 Methodology to Improve BWE and SC Metrics with Forecasting

This section utilized the methodology depicted in Figure 7.1, comprising four main stages: Demand Generation, Forecasting, Simulation, and Post-hoc Analysis.

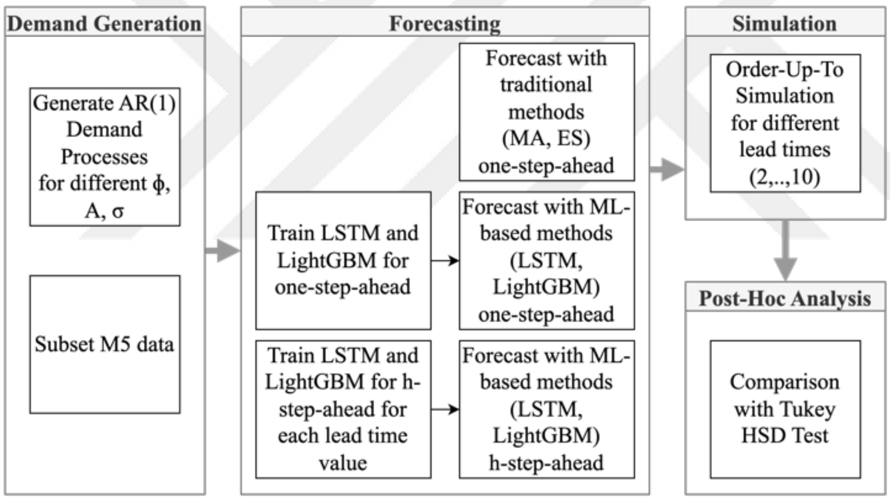

Figure 7.1: Methodology flow of Mitigation through Forecasting

In the first stage, demand data were generated using an autoregressive process of order 1, AR(1), to simulate different demand conditions. The autocorrelation (ϕ), seasonality amplitude (A), and variance (σ) were adjusted to explore a range of possible scenarios. In addition, a subset of the M5 competition dataset introduced a real-world dimension into the analysis.

In the second stage, the generated and real-world demand data are forecast using both the traditional and ML methods. The traditional methods MA and ES are employed for one-step-ahead forecasts, which predict the demand for one period in the future. Two ML models, LSTM and LightGBM, were trained and tested for both one-step-ahead and h-step-ahead forecasts. For h-step-ahead forecasting, multiple lead times are considered, enabling a more comprehensive assessment of the forecast horizon impact on supply chain performance.

In the third stage, the OUT method is simulated over lead times of two–ten periods to replicate the inventory control dynamics. The simulated inventory control system uses the forecasts generated in the previous step to replenish the stock based on the predicted future demand. The simulation aimed to replicate real-world supply chain dynamics and assess the performance of the forecasting models in terms of BWE, inventory level, inventory volatility, and service level. By adjusting the lead time, the simulation measures the effect of forecast accuracy over varying time horizons on the overall supply chain efficiency.

Finally, a post-hoc analysis was conducted using the Tukey Honest Significant Difference (HSD) test to statistically compare the performance of the different forecasting methods under each scenario. This analysis ensures that the differences observed in the supply chain metrics are statistically significant, providing a basis for evaluating the difference between one-step-ahead and h-step-ahead forecasting.

7.2 Order-Up-To-Level Inventory Control

A widely used inventory management approach, the OUT method, determines the optimal inventory level that minimizes costs while satisfying the demand Johnson and Thompson (1975). This method relies on forecasting future demand, determining the safety stock, and accounting for LT. This study implements a simulation of inventory control using the OUT under various LTs and different demand forecasts.

7.2.1 Lead Time Demand Calculation

The Lead Time Demand, D_{Lt}, is the total demand expected during lead time, LT. Depending on the approach used, this can be calculated by multiplying the forecasted demand F_t by L or summing the forecasted demand F_i, over the LT periods as follows:

$$D_{Lt} = \mathrm{F}_t * \mathrm{LT} \tag{20}$$

$$D_{Lt} = \sum_{i=t}^{t+L-1} \mathrm{F}_i \tag{21}$$

The multiplication approach can be used for all forecasting models, as it requires one-step-ahead forecasting, where the model predicts the value for the immediate next period. The forecast was based on past data up to the current period, and the prediction was made only for the next time point.

However, because these studies utilize MMSE and ARIMA forecasting with autoregressive processes to generate h-step-ahead forecasts, they are limited in addressing situations where demand parameters are unknown. ML-based methods, however, warrant attention owing to their flexibility in overcoming these barriers. By adapting their output shape, ML models can forecast multiple steps ahead without requiring prior knowledge of demand parameters or using recursive forecasting.

7.2.2 Safety Stock Calculation

Safety stock, SS, is an additional inventory buffer that accounts for demand variability during the LT. It is typically calculated based on service-level targets as follows:

$$SS = z \times \sigma_D \times \sqrt{\mathrm{LT}} \tag{22}$$

Equation (13) shows a typical SS level, where z is the safety factor (e.g., 1.96 for a 95% service level), σ_D is the standard deviation of demand, LT is the lead time. In this study, the safety stock was set to zero to directly measure the impact of LT Heydari *et al.* (2009).

7.2.3 Order-Up-To Level

The order-up-to level, S, determines the inventory position to which the inventory should be replenished and is calculated as follows:

$$S_t = D_{L,t} + SS \tag{23}$$

With zero safety stock, this simplifies to LT Demand.

7.2.4 Inventory Position

The inventory position $\mathrm{IP_t}$, is the inventory balance of the system that sums inventory level I_{t-1} at the end of the previous period, inventory currently in transit IT_t, and inventory AI_t that has arrived during period t. The equation is as follows:

$$\mathrm{IP}_t = I_{t-1} + \mathrm{IT}_t - \mathrm{D}_t + \mathrm{A}_t \tag{24}$$

7.2.5 Order Quantity

The order quantity Q_t, at time t is the amount of inventory required to raise the inventory position $\mathrm{IP_t}$, to the OUT, S:

$$Q_t = \max(0, S_t - \mathrm{IP}_t) \tag{25}$$

Equation (16) shows that Q_t is a non-negative variable that does not return excess inventory without a cost.

7.3 Supply Chain Metrics

7.3.1 Variance of Inventory

The variance of inventory, V_I quantifies the variability in inventory levels over time and can be represented by Equation (18) as follows:

$$V_\mathrm{I} = \frac{\mathrm{Var}(I_t)}{\mathrm{Var}(D_t)} \tag{26}$$

The variance of inventory is a key metric for evaluating the stability of inventory levels relative to demand fluctuations. A higher variance indicates a less stable system, leading to increased holding costs for more safety-stock needs.

7.3.2 Order Fulfillment Rate

The Order Fulfillment Rate (OFR) is the proportion of customer demand that is met without delay or backorder and can be expressed as follows:

$$B_t = max(0, D_t - I_t) \tag{27}$$

$$\mathrm{OFR_t} = \frac{D_t - B_t}{D_t} \tag{28}$$

If the backorder B_t is zero (19), then $\mathrm{OFR_t} = 1$, meaning that 100% of the demand is fulfilled. If backorders exist, $\mathrm{OFR_t}$ is less than 1 (20).

7.3.3 Average End Inventory

The average end inventory (22) refers to the average amount of inventory left at the end of each period, where $\mathrm{EI_t}$ is the inventory remaining at the end of the period and T is the total number of periods (21). The equations are expressed as follows:

$$\mathrm{EI_t} = max(0, \mathrm{EI_{t-1}} + \mathrm{AI_t} - D_t) \tag{29}$$

$$\text{Average EI} = \frac{\sum_{t=1}^{T} \text{End Inventory}_t}{T} \quad (30)$$

The average end inventory serves as an indicator of inventory efficiency, where higher values suggest overstocking and increased holding costs, whereas lower values may indicate a risk of stockouts and unmet demand. Maintaining an optimal average end inventory is crucial for balancing service levels with cost minimization in supply chain management.

7.4 Forecasting Results

7.4.1 Forecast Accuracy

Forecast accuracy was measured using the Root Mean Squared Error (RMSE) and Symmetric Mean Absolute Percentage Error (sMAPE). The RMSE is preferred for ranking forecast performance, for the AR(1) series, while sMAPE is preferred for M5 series, as it allows for comparison across different scales, ensuring consistent evaluation of diverse data.

7.4.1.1 Results for AR (1) Series without Seasonality

The forecasting methods ranked by RMSE show that LSTM and LightGBM consistently outperform ES and MA across different values of ϕ and σ. When analyzing sensitivity to ϕ and σ, it is evident that, as ϕ increases, the RMSE generally tends to increase for MA and ES, reflecting their difficulty in handling positively autocorrelated processes. Conversely, LSTM and LightGBM exhibited more resilience to changes in ϕ, with only slight increases in RMSE, maintaining lower errors even when ϕ approached 0.9.

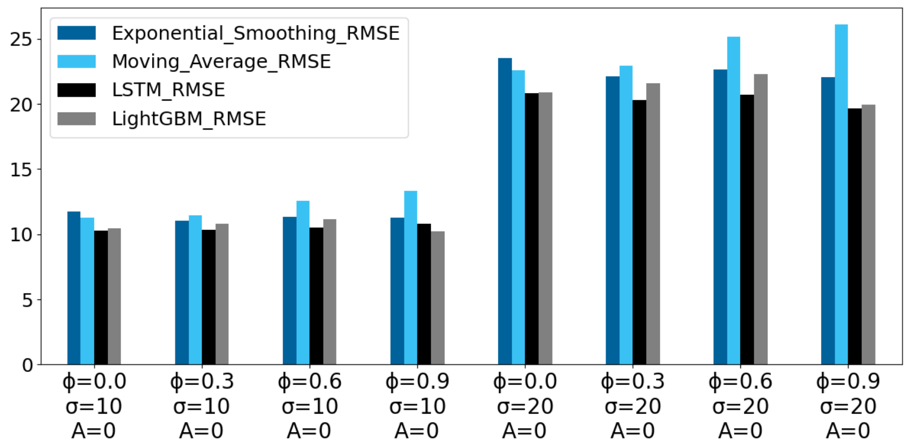

Figure 7.2: RMSE Comparison of AR (1) series with no seasonality.

Figure 7.2 shows that at $\phi = 0.0$ and $\sigma = 10$, the RMSE for LSTM is 10.26, which increases slightly to 10.83 at $\phi = 0.9$. LightGBM exhibits a similar trend, with RMSE values ranging from 10.44 at $\phi = 0.0$ to 10.21 at $\phi = 0.9$. In contrast, the RMSE for MA increased more significantly from 11.30 to 13.33 over the same range of ϕ, demonstrating its reduced effectiveness in handling higher autocorrelation.

As σ increased, the RMSE naturally grew rapidly for all methods, reflecting the increased variability in the demand process. For example, at $\phi = 0.0$ and $\sigma = 20$, the RMSE for LSTM increases to 20.82, whereas for LightGBM, it climbs to 20.88. Notably, ES and MA also showed substantial increases, with RMSE values doubling from their $\sigma = 10$ counterparts.

Overall, LSTM and LightGBM demonstrated superior performance, particularly in scenarios with higher autocorrelation and variability, making them preferable choices for forecasting in such contexts.

7.4.1.2 Results for AR (1) Series with Seasonality

The RMSE values in Figure 7.3 demonstrate that while ES and MA struggle with increasing seasonality amplitude and autocorrelation, LSTM and LightGBM maintain relatively lower error rates. As ϕ increases from 0.0 to 0.9, a notable increase in RMSE is observed for both ES and MA, particularly at higher amplitudes.

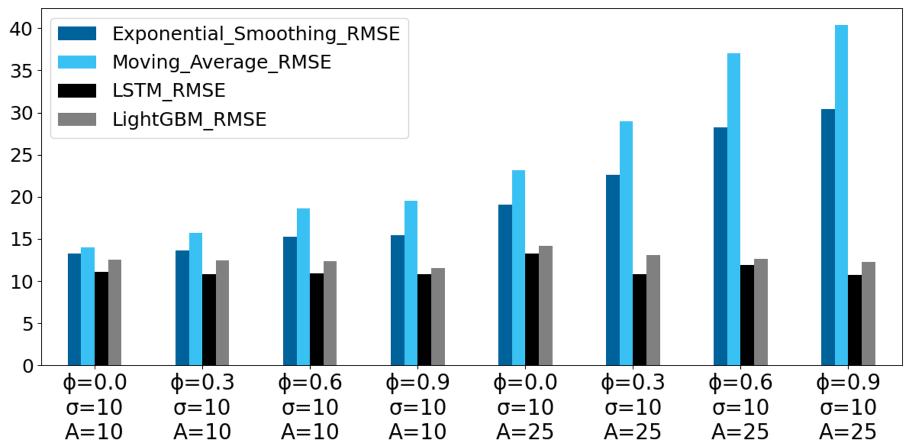

Figure 7.3: RMSE Comparison of AR (1) series with seasonality.

At $A = 25$ and $\phi = 0.9$, the RMSE for ES increased to 30.42, and for MA, it reached 40.36, reflecting their limitations in capturing complex seasonal patterns combined with positive autocorrelation. In contrast, LSTM and LightGBM display more stable RMSE values under the same conditions, with LSTM showing only a slight increase from 13.25 at $\phi = 0.0$ and $A = 25$ to 10.71 at $\phi = 0.9$ and $A = 25$.

The sensitivity analysis further underscores the impact of increasing the seasonal amplitude A. At $\phi = 0.0$, the RMSE for LSTM increases from 11.13 at $A = 10$ to 13.25 at $A = 25$, whereas the RMSE rises from 12.53 to 14.22 under the same conditions. This trend continued as ϕ increased, with all methods experiencing higher RMSE values at larger amplitudes. However, the performance of LSTM and LightGBM remained more consistent, indicating their robustness in dealing with seasonal demand patterns.

Overall, LSTM and LightGBM were better suited for forecasting in environments characterized by seasonality autocorrelation, as evidenced by their lower RMSE values and greater resilience to changes in ϕ and A.

Figure 7.4 illustrates the distribution of sMAPE values for each forecast method across different coefficients of variation (CV) groups for the M5 series. LightGBM consistently exhibited the lowest sMAPE values across all CV groups, with an overall sMAPE of 20.07, reflecting its robustness in accurately capturing demand patterns, even in highly variable datasets. Its performance is particularly strong in the lower CV group ($0.2 < CV \leq 0.3$), where the sMAPE remains below 19, indicating superior forecasting accuracy in more stable environments.

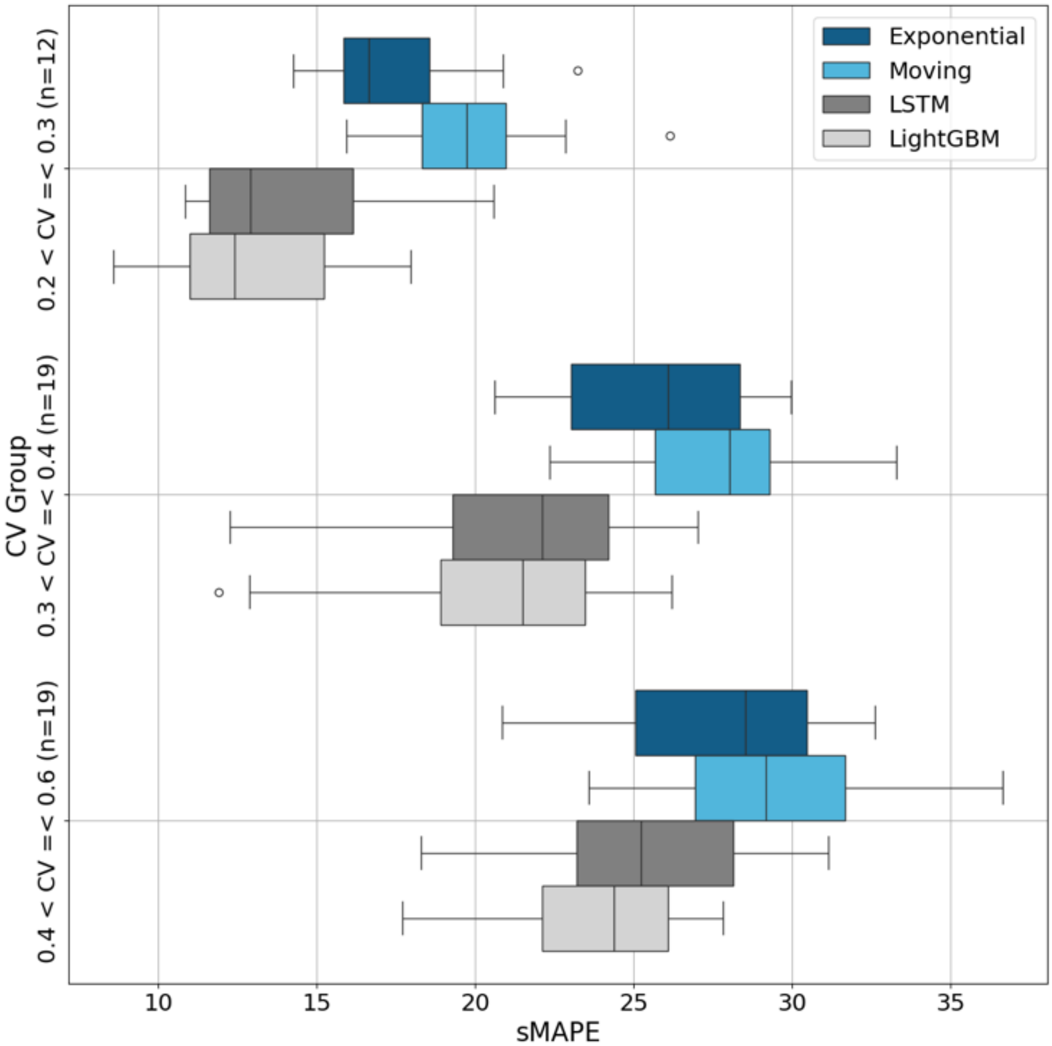

Figure 7.4: sMAPE Comparison of M5 competition series.

LSTM also performed well, with an overall sMAPE of 21.07. While slightly higher than LightGBM, LSTM shows strong adaptability, especially in the mid-range CV group ($0.3 < CV \leq 0.4$), where it manages to maintain a low sMAPE around 20.5, demonstrating its effectiveness in handling moderate variability.

On the other hand, ES and MA showed significantly higher sMAPEs, 24.49 and 26.60, respectively. These methods struggle more with datasets that exhibit higher variability, particularly in ($0.4 < CV \leq 0.6$) CV group (27. The pronounced gap in sMAPEs between MA, ES, and ML-based approaches is still evident in the higher CV groups.

7.5 Comparison Of One-Step and H-Step Ahead Forecasting

This section presents a comprehensive comparison of forecasting methods, forecasting horizons, various demand conditions, and their effects on critical supply chain metrics across both AR(1) demand processes and the M5 dataset. The performance of each method for the corresponding metric is illustrated with plots and supported by statistical significance.

Tables 7.1 and 7.3 show the simulation results for both the AR(1) series and M5 dataset, respectively, under varying conditions of autocorrelation (ϕ), seasonality (A), demand variance (σ), and coefficient of variation (CV). They compared the inventory management performance of the corresponding forecast methods at lead times 2 and 10. Tables 3.14 and 3.16 present the results of the Tukey Honest Significant Difference (HSD) test for the AR(1) and M5 datasets, providing statistical validation of the performance differences between the forecasting methods.

7.5.1.1 Amplification Ratio

The results indicate that, particularly at shorter lead times, the amplification ratios for LSTM and LightGBM were not significantly different from those of ES and MA.

As lead times increased, the differences between the ML-based methods and traditional methods became more apparent, although the improvements in the amplification ratio remained moderate. Specifically, both LSTM and LightGBM showed lower amplification ratios as lead times were extended, but the reduction was not as significant as might be expected given their superior forecast accuracy.

This suggests that while ML models provide better demand predictions, the traditional one-step-ahead structure limits their ability to fully mitigate BWE, which arises from forecast errors and demand variability amplification.

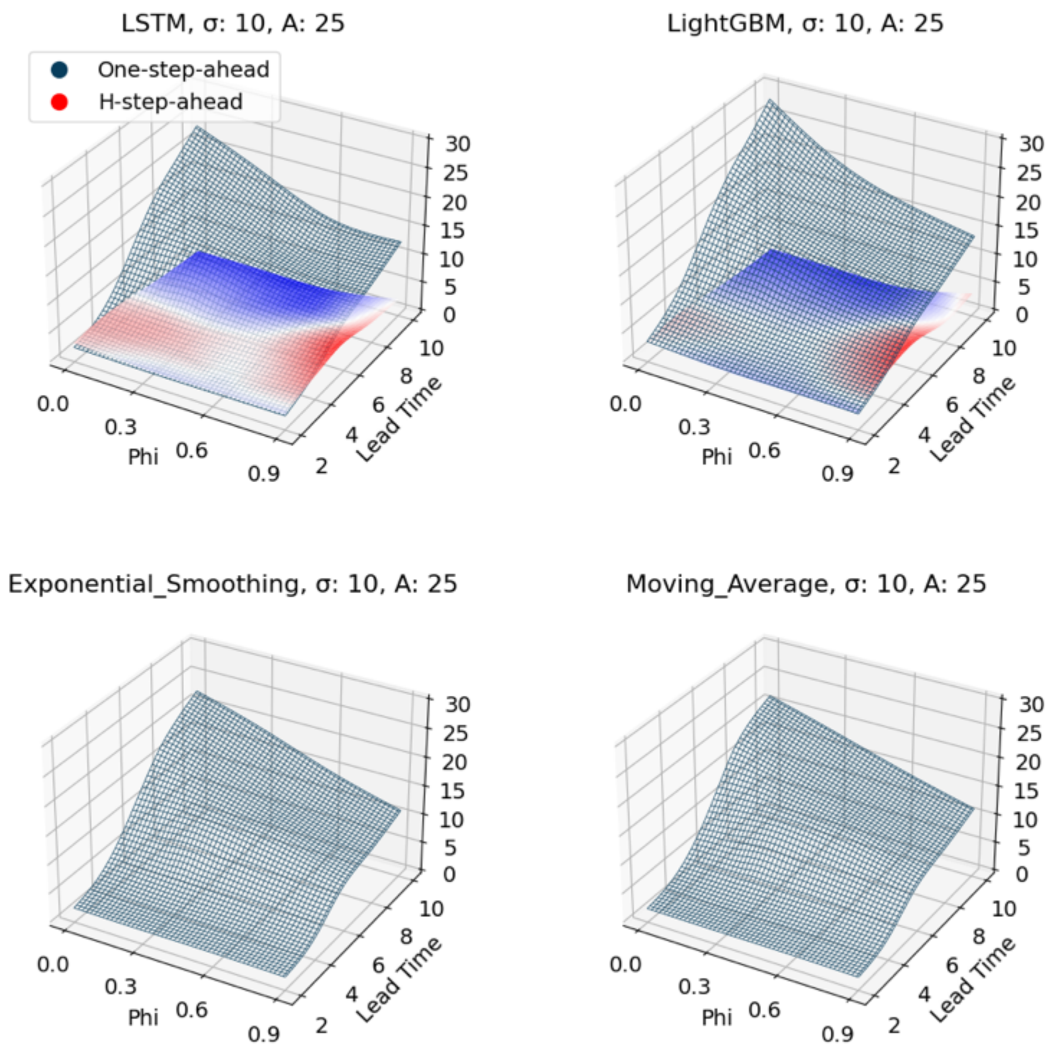

Figure 7.5: Amplification Ratio versus ϕ and LT for AR (1) Series.

The introduction of h-step-ahead forecasting for ML-based methods, specifically LSTM and LightGBM, led to substantial improvements in controlling the amplification ratio compared with both traditional methods and one-step-ahead ML models. Figure 7.5 and Figure. 7.6 show that, as lead times increased, the h-step-ahead forecasting methods consistently outperformed both the one-step-ahead approaches and traditional methods for both the AR(1) and M5 series.

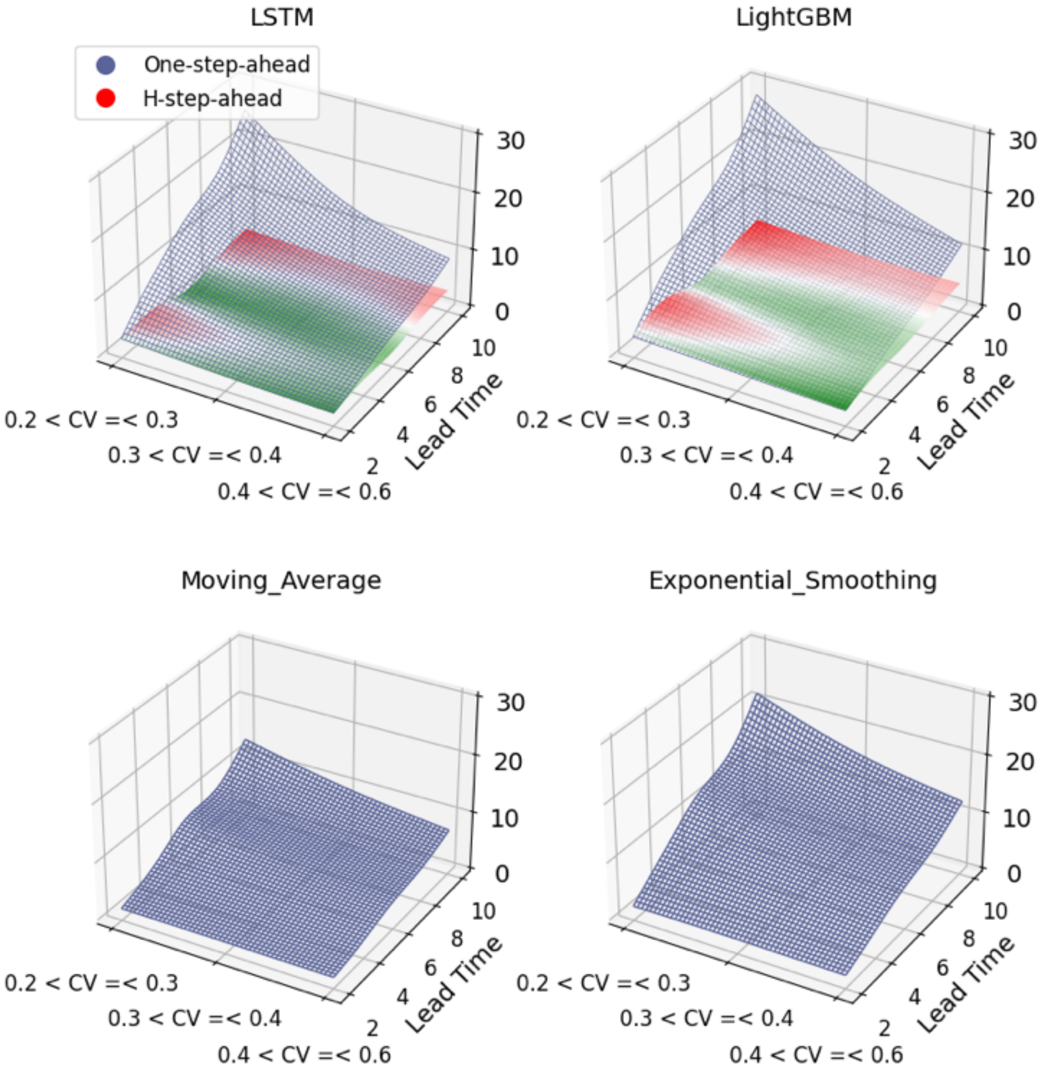

Figure 7.6: Amplification Ratio versus CV and LT for M5 Series.

The Tukey test results in Tables 7.2 and 7.4 further validate these findings, showing statistically significant improvements in the amplification ratio for h-step-ahead forecasts, particularly when compared to one-step-ahead forecasts from both the ML and traditional models. These differences are especially pronounced as lead times increase, where the benefit of multi-step forecasting becomes more evident in mitigating BWE.

7.5.2 Variance of Inventory

For one-step-ahead forecasts, the ML-based methods did not consistently exhibit a significantly lower variance of inventory compared to traditional methods. Although LSTM and LightGBM provided more accurate demand forecasts, this accuracy did not always translate into more stable inventory levels. This finding is more evident when lead times are short, where the variance of inventory remains comparable across both the traditional and ML methods.

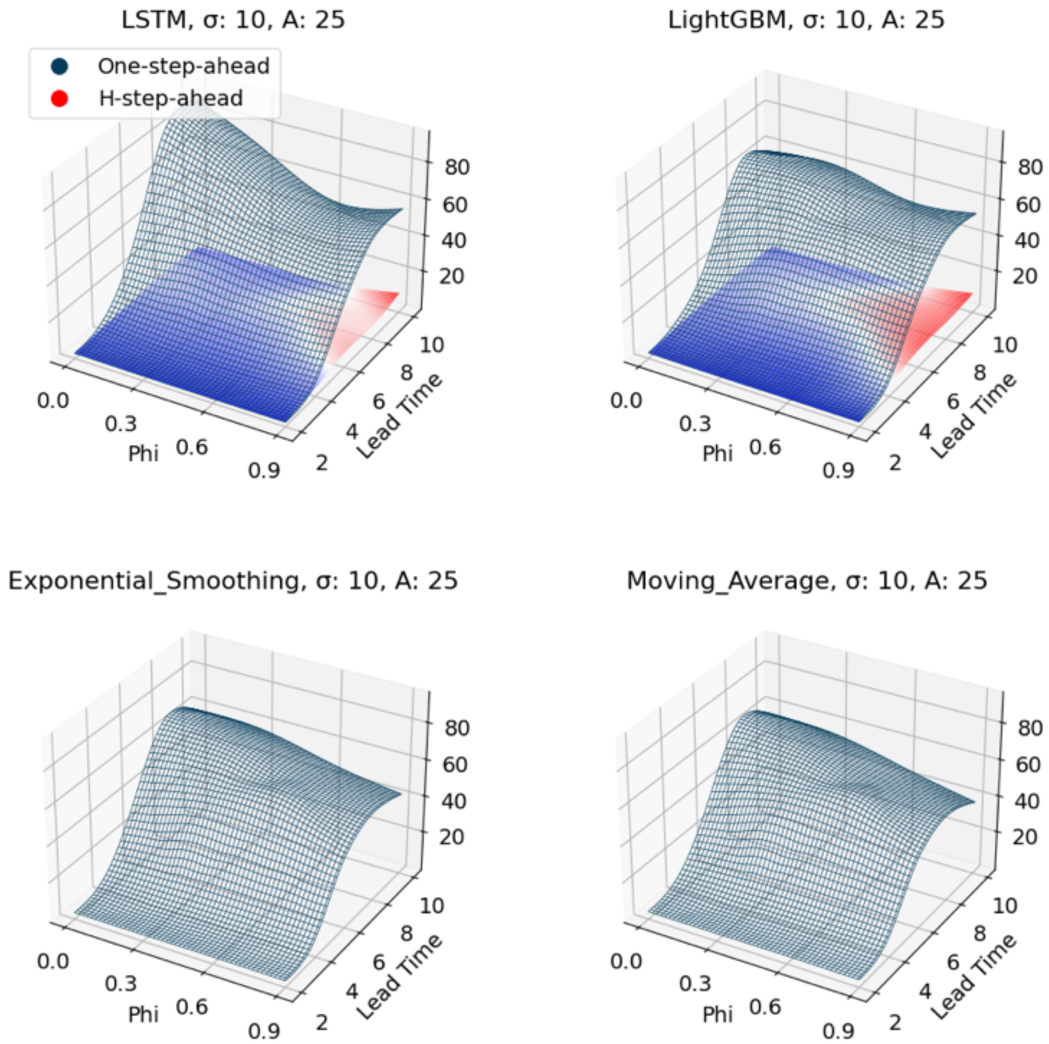

Figure 7.7: Variance of Inventory versus ϕ and LT for AR (1) Series.

As lead times increased, some differences began to emerge, with LSTM and LightGBM showing slightly lower inventory variance than ES and MA. However, these differences were not always statistically significant according to Tukey's test results. The results suggest that the forecast accuracy gains from ML methods in one-step-ahead forecasting are not fully realized in the context of inventory variance. The one-step-ahead nature of these forecasts limits their ability to anticipate long-term fluctuations in demand, which contributes to continued variability in inventory levels.

Figure 7.7 and Figure 7.8 show that the use of h-step-ahead forecasting with ML-based methods significantly reduced the variance of inventory compared to both traditional methods and one-step-ahead ML approaches.

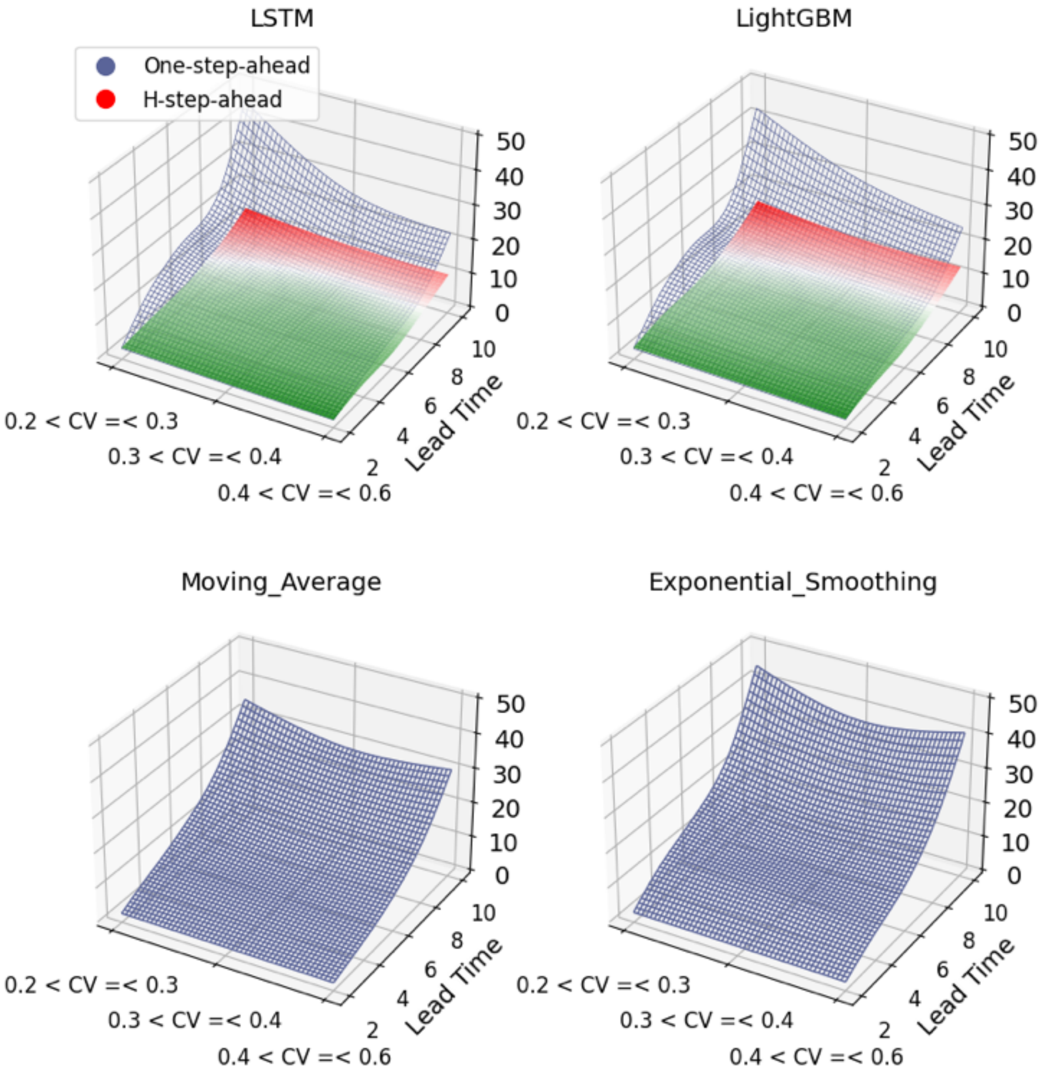

Figure 7.8: Variance of Inventory versus CV and LT for M5 Series.

The Tukey test results support these findings, showing statistically significant reductions in inventory variance when using h-step-ahead LSTM and LightGBM compared with both one-step-ahead forecasting methods and traditional approaches.

7.5.3 Order Fulfillment Rate

The results show that despite the superior forecast accuracy of LSTM and LightGBM in predicting demand, the OFRs did not significantly differ between ML-based methods and traditional methods. The Tukey test results indicate that while there are some statistically significant differences between methods for certain lead times, these differences are not as pronounced as expected, considering the forecast accuracy advantage of ML models.

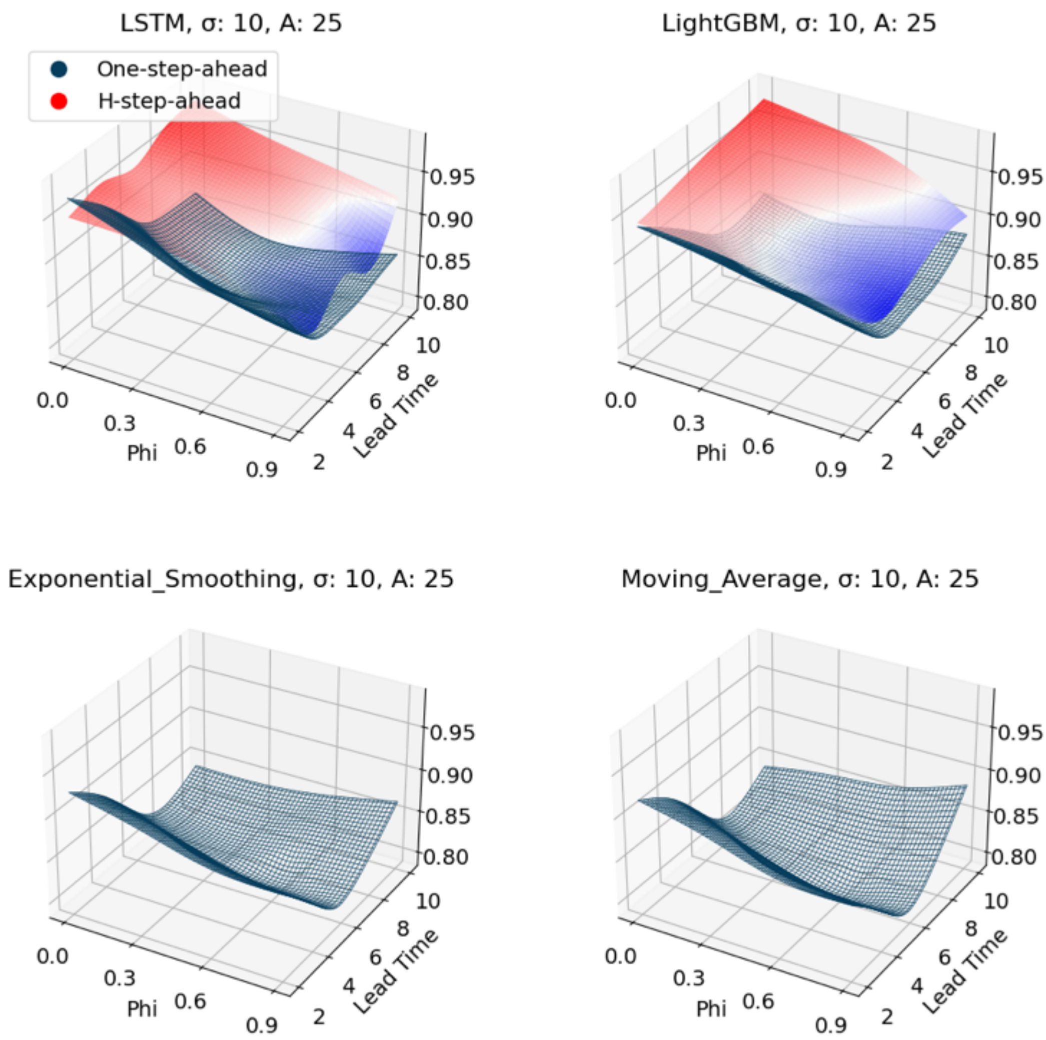

Figure 7.9: Average OFR versus ϕ and LT for AR (1) Series.

In concurrence with other metrics, Figure 7.9. And Figure 7.10. illustrate how h-step-ahead forecasting with LSTM and LightGBM offers a substantial improvement in OFR over one-step-ahead forecasting and traditional methods for both demand series.

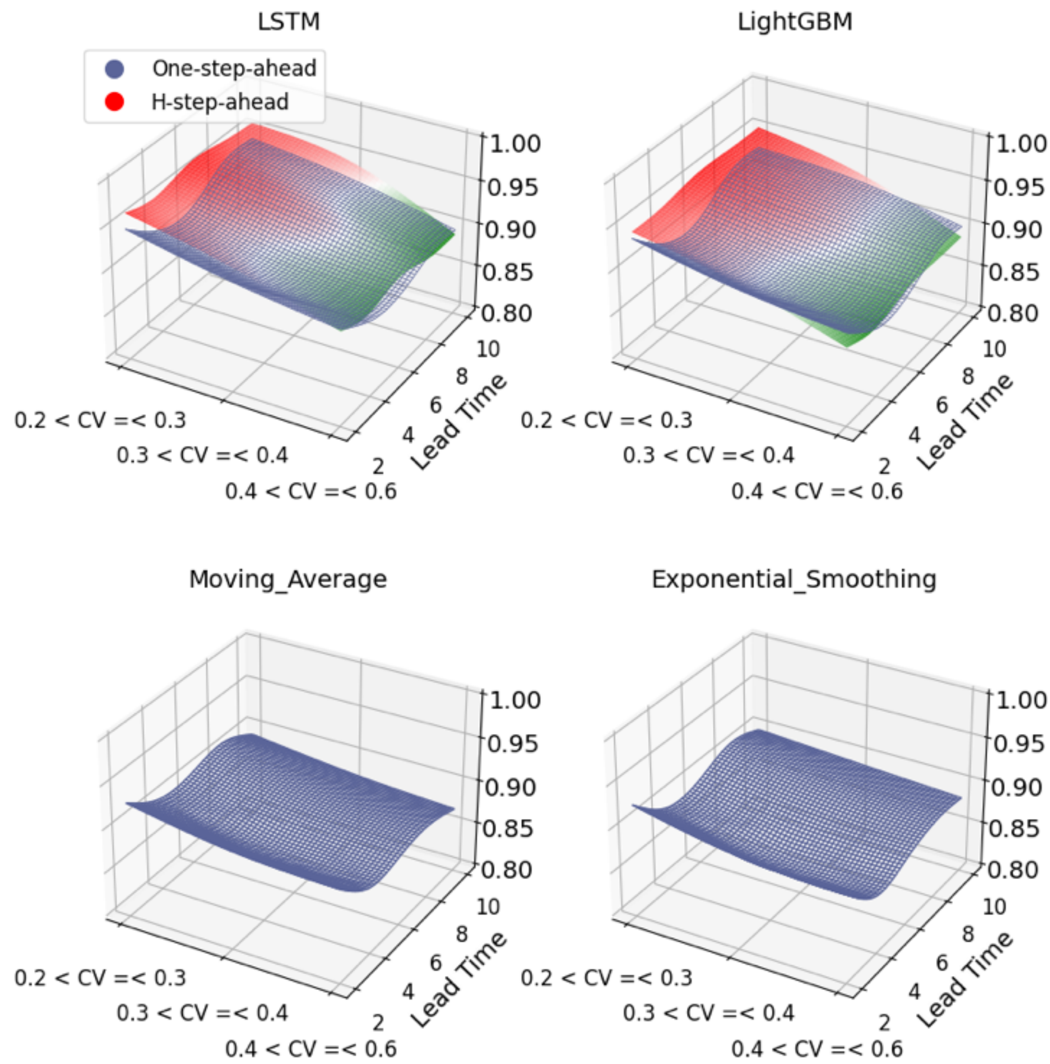

Figure 7.10: Average OFR versus CV and LT for M5 Series.

The Tukey test results indicate statistically significant differences in OFRs when comparing h-step-ahead ML methods with both one-step-ahead and traditional methods for the AR(1) series. These differences become more significant as the lead times increase. However, while there is a positive gap, the Tukey analysis for the M5 series does not confirm the statistical significance of the observed mean differences.

7.5.4 Average End Inventory

For the one-step-ahead forecasting comparison, Figure 7.11 and Figure 7.12 highlight that the one-step-ahead ML methods did not demonstrate a substantial improvement over traditional methods. As the lead-time increased, both categories of methods tended to converge in their performance, indicating that the superior forecast accuracy of LSTM and LightGBM did not translate into significant gains in inventory management.

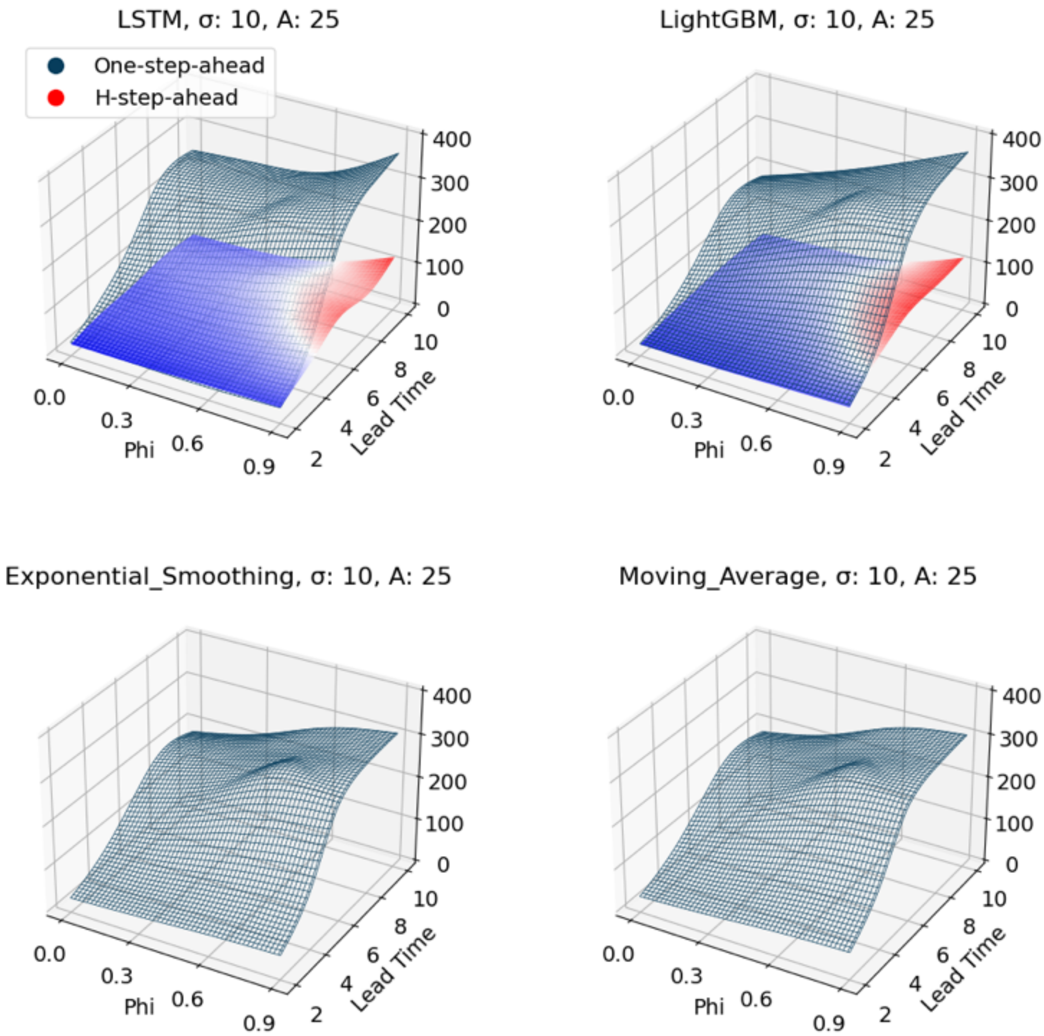

Figure 7.11: Average End Inventory versus ϕ and LT for AR (1) Series.

The introduction of h-step-ahead forecasting methods for ML-based models has yielded notable improvements in the Average End Inventory. This enhancement is particularly evident when compared with the one-step-ahead forecasts of traditional methods and ML-based approaches.

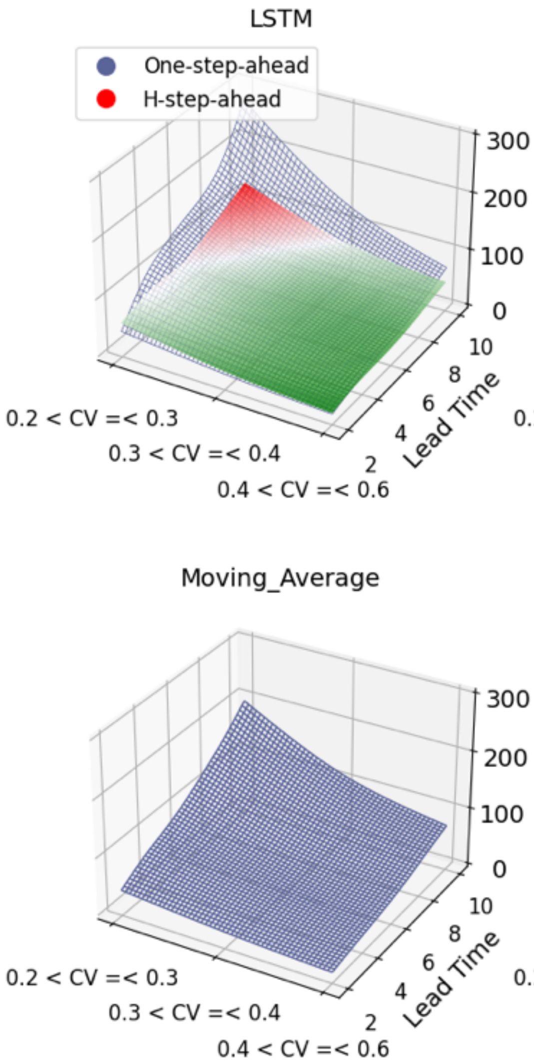

Figure 7.12: Average End Inventory versus CV and LT for M5 Series.

Tukey analysis shows a significant reduction in the Average End Inventory when utilizing h-step-ahead forecasting with LSTM and LightGBM, especially as lead times increase for the AR(1) series. However, the Tukey results for the M5 series do not support the significance of the mean differences, as is evident in Figure 7.12.

Table 7.1: Simulation Results of AR (1) Demand Process for $\sigma = 10$; LT = 2 / LT = 10

A	Method	Horizon	ϕ	AR	Avg. OFR	Var. of Inv.	Avg. Backorders	Avg. End Inv.
0.0	ES	One-step	0.0	2.21 / 33.04	0.97 / 0.94	1.75 / 36.16	3.66 / 6.40	11.14 / 70.23
			0.3	2.25 / 29.12	0.96 / 0.92	1.84 / 64.52	4.22 / 8.03	12.50 / 89.73
			0.6	1.76 / 23.53	0.96 / 0.91	1.44 / 71.55	4.68 / 9.94	13.71 / 108.91
			0.9	1.36 / 9.52	0.96 / 0.90	0.73 / 43.28	4.64 / 11.62	14.50 / 139.59
	LightGBM	H-step	0.0	2.02 / 2.70	0.97 / 0.98	1.32 / 6.77	3.34 / 2.30	10.29 / 27.34
			0.3	2.30 / 3.86	0.96 / 0.97	1.65 / 10.96	3.95 / 3.49	11.52 / 37.97
			0.6	2.51 / 6.30	0.95 / 0.94	1.45 / 18.90	5.49 / 6.71	13.44 / 54.07
			0.9	2.09 / 12.24	0.95 / 0.89	0.81 / 30.74	6.11 / 12.68	16.54 / 123.20
		One-step	0.0	2.24 / 20.42	0.97 / 0.96	1.45 / 13.74	3.48 / 4.11	10.80 / 47.07
			0.3	2.60 / 33.22	0.96 / 0.94	1.78 / 33.81	4.08 / 6.43	11.84 / 72.16
			0.6	2.72 / 28.69	0.95 / 0.91	1.56 / 43.85	5.51 / 9.43	14.08 / 97.19
			0.9	2.03 / 17.50	0.95 / 0.90	0.79 / 41.88	5.81 / 11.98	16.15 / 148.70
	LSTM	H-step	0.0	1.25 / 1.35	0.98 / 0.98	1.22 / 5.21	2.41 / 1.94	10.21 / 22.62
			0.3	1.53 / 1.94	0.97 / 0.97	1.30 / 8.45	3.76 / 3.62	10.02 / 28.00
			0.6	1.69 / 3.90	0.96 / 0.94	1.16 / 13.05	4.56 / 6.40	12.35 / 41.18
			0.9	1.84 / 8.98	0.96 / 0.91	0.83 / 25.08	4.61 / 11.29	17.66 / 107.52
		One-step	0.0	1.26 / 1.70	0.98 / 0.99	1.24 / 6.40	2.34 / 1.11	10.45 / 28.86
			0.3	1.67 / 12.12	0.96 / 0.94	1.21 / 18.85	4.53 / 6.24	9.10 / 46.72
			0.6	1.92 / 21.66	0.96 / 0.92	1.19 / 38.48	5.07 / 8.92	12.06 / 84.29
			0.9	1.65 / 13.13	0.94 / 0.88	0.55 / 36.63	7.07 / 14.35	12.04 / 132.02
	MA	One-step	0.0	1.39 / 13.90	0.97 / 0.95	1.33 / 24.75	3.27 / 4.91	9.96 / 53.36
			0.3	1.59 / 16.94	0.96 / 0.93	1.60 / 58.91	3.94 / 7.51	11.67 / 82.91
			0.6	1.37 / 16.20	0.96 / 0.91	1.48 / 68.62	4.53 / 9.53	13.24 / 102.41
			0.9	1.29 / 7.94	0.96 / 0.90	0.82 / 45.30	4.88 / 11.64	15.24 / 136.24
10.0	ES	One-step	0.0	2.08 / 30.86	0.96 / 0.90	1.78 / 58.26	4.60 / 9.49	14.17 / 108.81
			0.3	1.96 / 28.45	0.95 / 0.88	1.78 / 83.52	5.76 / 11.02	17.34 / 136.60
			0.6	1.86 / 24.33	0.94 / 0.86	1.91 / 94.46	7.00 / 13.68	20.99 / 181.15
			0.9	1.53 / 12.15	0.94 / 0.87	1.17 / 54.85	7.06 / 13.15	21.94 / 188.75
	LightGBM	H-step	0.0	2.30 / 1.98	0.96 / 0.98	1.61 / 4.76	3.99 / 2.30	13.40 / 29.18
			0.3	2.83 / 2.23	0.95 / 0.97	1.66 / 6.51	5.52 / 3.28	16.12 / 35.73
			0.6	3.19 / 3.10	0.94 / 0.95	1.55 / 7.99	6.76 / 6.03	18.04 / 54.17
			0.9	2.53 / 9.91	0.93 / 0.88	0.96 / 24.25	8.04 / 13.79	20.04 / 119.39
		One-step	0.0	2.69 / 30.17	0.96 / 0.93	1.70 / 31.44	4.13 / 6.63	13.74 / 81.39
			0.3	3.16 / 34.96	0.95 / 0.89	1.76 / 59.43	5.81 / 9.91	16.90 / 124.30
			0.6	3.28 / 34.08	0.94 / 0.87	1.66 / 76.04	7.00 / 13.24	19.18 / 178.35
			0.9	2.43 / 21.64	0.93 / 0.87	0.99 / 65.53	7.68 / 13.61	20.78 / 213.74
	LSTM	H-step	0.0	1.54 / 1.21	0.97 / 0.98	1.25 / 4.50	3.68 / 2.63	11.81 / 27.63
			0.3	2.10 / 1.13	0.96 / 0.96	1.31 / 4.90	5.04 / 4.00	14.36 / 26.46
			0.6	2.62 / 1.84	0.95 / 0.94	1.36 / 6.89	6.16 / 6.66	17.76 / 46.21
			0.9	2.34 / 6.61	0.94 / 0.92	0.97 / 22.35	7.04 / 9.70	21.55 / 105.98
		One-step	0.0	1.31 / 15.11	0.98 / 0.93	1.13 / 44.67	2.69 / 6.96	12.80 / 98.03
			0.3	1.76 / 23.14	0.96 / 0.87	1.06 / 63.74	4.75 / 12.02	13.08 / 137.46
			0.6	2.17 / 26.53	0.96 / 0.87	1.33 / 79.23	5.01 / 12.94	18.29 / 187.68
			0.9	1.93 / 17.66	0.93 / 0.86	0.76 / 56.44	7.39 / 14.84	17.31 / 201.30
	MA	One-step	0.0	1.58 / 22.69	0.96 / 0.90	1.84 / 55.68	4.82 / 9.05	14.80 / 101.20
			0.3	1.65 / 23.64	0.95 / 0.88	2.10 / 78.16	5.92 / 10.38	17.88 / 128.04
			0.6	1.69 / 22.24	0.94 / 0.86	2.30 / 89.86	7.53 / 13.31	22.55 / 171.13
			0.9	1.49 / 11.80	0.93 / 0.87	1.49 / 53.57	7.83 / 13.38	24.23 / 189.89
25.0	ES	One-step	0.0	1.93 / 22.48	0.93 / 0.84	1.95 / 58.27	8.42 / 12.39	25.89 / 182.21
			0.3	1.85 / 18.82	0.91 / 0.84	1.95 / 57.24	10.96 / 11.89	33.48 / 218.60
			0.6	1.86 / 14.58	0.89 / 0.85	2.09 / 48.73	14.15 / 10.96	42.95 / 277.26
			0.9	1.70 / 11.17	0.88 / 0.87	1.83 / 42.70	15.17 / 10.92	48.71 / 308.38
	LightGBM	H-step	0.0	2.90 / 0.92	0.94 / 0.98	1.31 / 2.51	6.30 / 2.86	19.11 / 37.43
			0.3	2.24 / 0.79	0.94 / 0.96	0.88 / 3.05	6.91 / 4.59	20.61 / 47.22
			0.6	2.17 / 0.86	0.92 / 0.94	0.80 / 3.94	8.44 / 7.06	24.97 / 62.63
			0.9	2.14 / 3.39	0.91 / 0.90	0.70 / 9.14	10.55 / 12.62	29.23 / 118.50
		One-step	0.0	3.01 / 28.01	0.94 / 0.86	1.44 / 55.49	6.83 / 11.41	20.84 / 179.97
			0.3	2.47 / 20.37	0.93 / 0.85	1.09 / 59.17	8.02 / 11.58	23.60 / 233.11
			0.6	2.34 / 16.09	0.91 / 0.87	1.07 / 50.64	9.80 / 10.15	30.00 / 295.32
			0.9	2.19 / 13.62	0.90 / 0.88	0.89 / 53.65	11.36 / 9.86	34.34 / 364.93
	LSTM	H-step	0.0	2.74 / 0.54	0.95 / 0.97	1.19 / 2.26	5.68 / 2.96	19.67 / 35.72
			0.3	2.10 / 0.49	0.95 / 0.95	0.84 / 2.74	6.11 / 5.96	21.24 / 40.52
			0.6	2.13 / 0.48	0.94 / 0.94	0.76 / 3.50	7.34 / 8.22	25.65 / 62.24
			0.9	2.07 / 2.40	0.91 / 0.92	0.64 / 9.33	10.81 / 10.48	28.41 / 120.00
		One-step	0.0	1.96 / 23.31	0.97 / 0.86	1.20 / 86.08	3.09 / 11.28	24.30 / 242.06
			0.3	1.79 / 18.19	0.94 / 0.83	0.82 / 68.36	6.34 / 13.17	21.94 / 252.52
			0.6	1.79 / 13.23	0.91 / 0.84	0.63 / 50.37	9.29 / 12.84	21.47 / 275.09
			0.9	1.84 / 12.65	0.91 / 0.86	0.69 / 56.20	9.98 / 12.35	28.41 / 362.27
	MA	One-step	0.0	1.82 / 21.60	0.92 / 0.84	2.46 / 55.88	9.65 / 12.66	29.58 / 181.79
			0.3	1.79 / 18.20	0.90 / 0.85	2.55 / 54.47	12.40 / 11.36	37.80 / 216.77
			0.6	1.81 / 14.64	0.88 / 0.86	2.70 / 46.19	16.22 / 10.60	49.15 / 277.36
			0.9	1.67 / 11.50	0.87 / 0.89	2.38 / 38.24	17.33 / 10.22	55.40 / 304.36

Table 7.2: Tukey HSD Test Results Of Ar (1) For Lt = 2 / Lt = 10

		Amplification Ratio		Avg. End Inventory		Avg. OFR		Variance of Inventory	
Group 1	Group 2	p-adj[a]	Mean diff.	p-adj	Mean diff.	p-adj	Mean diff.	p-adj	Mean diff.
ES - One-step	LightGBM - H-step	0.00 / 0.00	0.50 / -15.84	0.80 / 0.00	-4.05 / -99.82	0.99 / 0.00	0.00 / 0.06	0.08 / 0.00	-0.38 / -44.60
	LightGBM - One-step	0.00 / 0.85	0.67 / 2.51	0.96 / 1.00	-2.66 / -3.90	1.00 / 0.89	0.00 / 0.01	0.42 / 0.28	-0.26 / -12.33
	LSTM - H-step	1.00 / 0.00	0.01 / -17.31	0.74 / 0.00	-4.34 / -109.37	0.59 / 0.00	0.01 / 0.06	0.00 / 0.00	-0.54 / -46.72
	LSTM - One-step	0.88 / 0.15	-0.13 / -5.32	0.64 / 1.00	-4.89 / -7.48	0.75 / 0.92	0.01 / 0.01	0.00 / 0.33	-0.58 / -11.83
	MA - One-step	0.08 / 0.23	-0.32 / -4.85	1.00 / 1.00	1.31 / -6.87	1.00 / 1.00	0.00 / 0.00	0.90 / 0.99	0.14 / -3.09
LightGBM - H-step	LightGBM - One-step	0.71 / 0.00	0.17 / 18.36	1.00 / 0.01	1.39 / 95.92	1.00 / 0.02	0.00 / -0.04	0.96 / 0.00	0.12 / 32.27
	MA - One-step	0.00 / 0.00	-0.82 / 11.00	0.54 / 0.01	5.36 / 92.94	0.97 / 0.00	-0.01 / -0.05	0.00 / 0.00	0.52 / 41.51
LightGBM - One-step	MA - One-step	0.00 / 0.01	-0.99 / -7.36	0.81 / 1.00	3.97 / -2.98	1.00 / 0.98	0.00 / -0.01	0.05 / 0.60	0.40 / 9.24
LSTM - H-step	LightGBM - H-step	0.00 / 0.98	0.49 / 1.46	1.00 / 1.00	0.30 / 9.55	0.89 / 1.00	-0.01 / 0.00	0.84 / 1.00	0.16 / 2.12
	LightGBM - One-step	0.00 / 0.00	0.66 / 19.82	0.99 / 0.00	1.68 / 105.47	0.73 / 0.01	-0.01 / -0.05	0.33 / 0.00	0.28 / 34.39
	LSTM - One-step	0.83 / 0.00	-0.14 / 11.98	1.00 / 0.00	-0.55 / 101.89	1.00 / 0.01	0.00 / -0.05	1.00 / 0.00	-0.04 / 34.89
	MA - One-step	0.06 / 0.00	-0.34 / 12.46	0.48 / 0.00	5.66 / 102.49	0.42 / 0.00	-0.02 / -0.06	0.00 / 0.00	0.68 / 43.63
LSTM - One-step	LightGBM - H-step	0.00 / 0.00	0.63 / -10.52	1.00 / 0.01	0.84 / -92.34	0.96 / 0.01	-0.01 / 0.04	0.68 / 0.00	0.20 / -32.77
	LightGBM - One-step	0.00 / 0.01	0.80 / 7.83	0.98 / 1.00	2.23 / 3.58	0.86 / 1.00	-0.01 / 0.00	0.19 / 1.00	0.32 / -0.50
	MA - One-step	0.59 / 1.00	-0.19 / 0.48	0.37 / 1.00	6.20 / 0.60	0.59 / 0.99	-0.01 / -0.01	0.00 / 0.66	0.72 / 8.74

Table 7.3: Simulation Results of M5 for LT = 2 / LT = 10

Method	**Horizon**	**CV Group**	**AR**	**Avg. OFR**	**Var. of Inv.**	**Avg. Backorders**	**Avg. End Inv.**
ES	One-step	0.2 < CV =< 0.3	1.92 / 21.55	0.93 / 0.90	1.90 / 44.49	12.61 / 15.35	38.80 / 222.04
		0.3 < CV =< 0.4	1.94 / 15.70	0.90 / 0.89	1.78 / 36.48	7.70 / 7.35	23.68 / 131.22
		0.4 < CV =< 0.6	1.74 / 12.37	0.90 / 0.88	1.55 / 40.92	4.55 / 4.82	13.86 / 86.38
LightGBM	H-step	0.2 < CV =< 0.3	3.58 / 5.48	0.94 / 0.95	2.39 / 15.31	9.62 / 7.96	44.69 / 130.13
		0.3 < CV =< 0.4	2.80 / 4.50	0.92 / 0.93	2.01 / 12.63	6.66 / 5.40	24.81 / 68.80
		0.4 < CV =< 0.6	2.11 / 4.48	0.89 / 0.89	1.42 / 12.78	5.09 / 4.67	13.99 / 49.27
	One-step	0.2 < CV =< 0.3	3.12 / 27.88	0.94 / 0.93	2.05 / 42.83	10.74 / 12.09	36.46 / 254.57
		0.3 < CV =< 0.4	2.57 / 17.86	0.91 / 0.92	1.74 / 30.13	6.66 / 5.68	21.73 / 140.67
		0.4 < CV =< 0.6	2.08 / 11.42	0.90 / 0.90	1.35 / 24.28	4.40 / 4.48	13.00 / 79.19
LSTM	H-step	0.2 < CV =< 0.3	2.56 / 3.82	0.96 / 0.96	2.37 / 13.08	6.43 / 7.08	53.72 / 123.28
		0.3 < CV =< 0.4	1.92 / 3.12	0.94 / 0.94	1.72 / 10.81	5.29 / 5.00	26.41 / 66.65
		0.4 < CV =< 0.6	1.58 / 3.18	0.90 / 0.89	1.24 / 10.47	4.34 / 4.56	14.11 / 46.12
	One-step	0.2 < CV =< 0.3	2.77 / 25.23	0.95 / 0.94	2.13 / 45.06	9.30 / 10.68	39.11 / 270.73
		0.3 < CV =< 0.4	2.06 / 13.59	0.92 / 0.92	1.57 / 26.93	6.42 / 5.62	21.73 / 138.21
		0.4 < CV =< 0.6	1.66 / 8.85	0.91 / 0.90	1.24 / 22.61	4.28 / 4.56	12.44 / 72.20
MA	One-step	0.2 < CV =< 0.3	1.49 / 13.34	0.93 / 0.89	1.57 / 34.60	12.65 / 15.66	38.84 / 196.21
		0.3 < CV =< 0.4	1.47 / 9.62	0.91 / 0.88	1.48 / 28.46	7.64 / 8.01	23.39 / 111.52
		0.4 < CV =< 0.6	1.38 / 7.26	0.90 / 0.87	1.30 / 30.34	4.44 / 5.19	13.36 / 73.18

Table 7.4: Tukey HSD Test Results of M5 for LT = 2 / LT = 10

		Amplification Ratio		Avg. End Inventory		Avg. OFR		Variance of Inventory	
Group 1	Group 2	p-adj[a]	Mean diff.	p-adj	Mean diff.	p-adj	Mean diff.	p-adj	Mean diff.
ES - One-step	LightGBM - H-step	0.00 / 0.00	-0.86 / 11.12	1.00 / 0.25	-1.89 / -59.88	0.78 / 0.00	-0.01 / 0.03	0.70 / 0.00	-0.16 / -26.76
	LightGBM - One-step	0.00 / 0.41	-0.66 / -1.98	1.00 / 1.00	-1.63 / 8.67	0.53 / 0.00	0.01 / 0.02	1.00 / 0.00	-0.06 / -9.14
	LSTM - H-step	0.98 / 0.00	-0.08 / 12.53	0.96 / 0.19	4.72 / -63.54	0.00 / 0.00	0.03 / 0.04	1.00 / 0.00	-0.03 / -28.87
	LSTM - One-step	0.43 / 0.83	-0.22 / 1.26	1.00 / 1.00	-1.20 / 8.96	0.02 / 0.00	0.02 / 0.03	0.78 / 0.00	-0.14 / -10.45
	MA - One-step	0.01 / 0.00	0.42 / 6.23	1.00 / 0.98	-0.29 / -18.70	0.94 / 0.35	0.00 / -0.01	0.10 / 0.00	-0.29 / -9.44
LightGBM - H-step	LightGBM - One-step	0.49 / 0.00	0.21 / -13.09	0.99 / 0.13	-3.52 / 68.54	1.00 / 0.95	0.00 / 0.00	0.36 / 0.00	-0.22 / 17.62
	MA - One-step	0.00 / 0.00	1.29 / -4.89	1.00 / 0.66	-2.18 / 41.18	1.00 / 0.00	0.00 / -0.04	0.00 / 0.00	-0.44 / 17.32
LightGBM - One-step	MA - One-step	0.00 / 0.00	1.08 / 8.20	1.00 / 0.92	1.34 / -27.37	0.97 / 0.00	0.00 / -0.04	0.29 / 1.00	-0.23 / -0.30
LSTM - H-step	LightGBM - H-step	0.00 / 0.75	-0.78 / -1.41	1.00 / 1.00	-2.82 / 3.66	0.00 / 0.80	-0.02 / -0.01	0.55 / 0.91	0.18 / 2.11
	LightGBM - One-step	0.00 / 0.00	-0.57 / -14.50	0.86 / 0.09	-6.34 / 72.20	0.01 / 0.25	-0.02 / -0.01	1.00 / 0.00	-0.03 / 19.73
	LSTM - One-step	0.87 / 0.00	-0.13 / -11.27	0.89 / 0.09	-5.92 / 72.49	0.33 / 0.43	-0.01 / -0.01	0.89 / 0.00	-0.12 / 18.42
	MA - One-step	0.00 / 0.00	0.51 / -6.30	0.94 / 0.58	-5.01 / 44.84	0.00 / 0.00	-0.02 / -0.05	0.16 / 0.00	-0.26 / 19.43
LSTM - One-step	LightGBM - H-step	0.00 / 0.00	-0.64 / 9.86	0.99 / 0.12	3.09 / -68.83	0.39 / 0.99	-0.01 / 0.00	0.07 / 0.00	0.30 / -16.31
	LightGBM - One-step	0.00 / 0.03	-0.44 / -3.23	1.00 / 1.00	-0.43 / -0.29	0.65 / 1.00	-0.01 / 0.00	0.97 / 0.99	0.09 / 1.31
	MA - One-step	0.00 / 0.00	0.64 / 4.97	1.00 / 0.91	0.91 / -27.66	0.20 / 0.00	-0.01 / -0.04	0.78 / 1.00	-0.14 / 1.01

[a] p-adj values below 0.05 denotes statistically significant difference.

The findings from this study underscore the critical role of advanced forecasting methods in mitigating BWE and improving overall supply chain efficiency. By leveraging machine learning-based forecasting techniques such as Long Short-Term Memory (LSTM) and LightGBM, this research demonstrates that multi-step forecasting significantly outperforms traditional approaches in stabilizing inventory levels, reducing demand amplification, and enhancing order fulfillment rates. The empirical results highlight that an improved forecasting strategy does not merely reduce variance but also contributes to better decision-making in inventory control and production planning. Furthermore, the comparison between one-step and h-step-ahead forecasts reveals that long-term forecasting approaches can enhance supply chain resilience by enabling proactive adjustments to fluctuations in demand. These insights provide a strong foundation for the integration of AI-driven forecasting tools in real-world supply chain operations. Future research should explore the scalability of these models across various industries and assess their adaptability to different market dynamics and lead time structures. Ultimately, the adoption of advanced forecasting methods presents a strategic opportunity for firms seeking to minimize uncertainty, enhance operational stability, and build more responsive and agile supply chains.

8. CONCLUSION

This research is structured around three core empirical analyses and a complementary study on advanced forecasting methods, each focusing on different dimensions of the BWE. The findings provide novel insights into the industry-specific nature of BWE, the role of importation in demand amplification, the impact of global disruptions such as the COVID-19 pandemic, and the effectiveness of advanced forecasting techniques in mitigating these effects.

The first study employs industry-level data to investigate the variability of BWE across different sectors. The findings challenge conventional assumptions about the uniformity of BWE drivers. Notably, while job losses demonstrate a strong relationship with BWE, other commonly presumed factors, such as inventory smoothing, backorders, and capacity utilization, exhibit inconsistent or even mitigating effects in certain industries. These results underscore the need for industry-specific intervention strategies rather than generic BWE mitigation approaches. By identifying sectors where BWE is more pronounced and those where it is naturally dampened, the study provides a foundation for best-practice learning across industries.

The second study examines the influence of import dependence on demand variability and its relationship with BWE. While longer lead times have traditionally been considered a primary driver of BWE, the results indicate that lead time alone is not a direct contributor. Instead, the key factor exacerbating demand amplification is the lack of supply chain visibility and inadequate information-sharing mechanisms in globally dispersed supply chains. This study highlights that firms should prioritize collaboration and real-time data exchange rather than solely focusing on nearshoring or lead-time reduction as a mitigation strategy. These insights are particularly valuable for industries with high reliance on imported intermediate goods, offering targeted solutions for managing demand fluctuations.

The third study assesses the effects of the COVID-19 pandemic on BWE at the industry level, providing empirical evidence on how supply chains responded to extreme demand and supply shocks. The findings confirm that certain industries experienced a significant surge in BWE, particularly those with high exposure to demand volatility and global sourcing constraints. However, the results also show that some sectors were more resilient, exhibiting lower demand amplification due to effective risk management and supply chain agility. This study highlights the need for industries to develop adaptive planning strategies and crisis response mechanisms to mitigate BWE during future global disruptions.

The final study explores the effectiveness of machine learning-based forecasting methods—specifically Long Short-Term Memory (LSTM) and LightGBM—in reducing demand amplification and improving supply chain performance. Unlike traditional forecasting techniques, these models demonstrate superior predictive accuracy in handling demand fluctuations, leading to reduced inventory variance and higher order fulfillment rates. The study provides empirical support for integrating advanced forecasting methodologies into supply chain decision-making processes to enhance overall resilience.

Additionally, the dissertation presents a causal framework mapping the interrelationships between BWE drivers and mitigation strategies. This framework highlights the critical role of information sharing and collaborative planning in reducing demand variability. The findings contribute to the broader discourse on supply chain localization, suggesting that while local sourcing strategies may help mitigate some risks, the key to long-term resilience lies in improving data-driven decision-making, transparency, and supply chain coordination.

8.1 Limitations

This study does not account for the potential occurrence of BWE at the information flow level. While the analysis primarily focuses on demand amplification in material flows, distortions arising from delayed information processing, prolonged backorder queues, or frequent order cancellations are beyond the scope of this research. Future studies could investigate these aspects to provide a more holistic perspective on BWE dynamics across different supply chain flows.

The empirical analysis is limited to manufacturing industries, utilizing industry-level panel data to examine BWE. Although this approach enables a structured and detailed assessment, it may constrain the generalizability of findings beyond the manufacturing sector. The exclusion of retailers and wholesalers, which operate under distinct economic dynamics due to their proximity to end customers, represents a limitation. Given that the transmission of demand variability differs between upstream and downstream supply chain actors, future research could extend the analysis to these sectors to gain a more comprehensive understanding of BWE propagation.

Despite the advantages of industry-level data in capturing broader trends, certain methodological constraints exist. The study relies on the amplification ratio as the primary metric for measuring BWE, which reflects the net effect of both amplifying and attenuating forces within the supply chain. However, this measure does not differentiate between cases where variance amplification is present but remains below a threshold that would conventionally define the presence of BWE. The assumption that positive changes in the amplification ratio, even if below one, indicate increased variance poses a potential limitation in interpreting results. Additionally, while industry-level data allows for macro-level insights, it may obscure firm-level decision-making factors that contribute to BWE. Future research could address this limitation by incorporating firm-specific data or integrating qualitative assessments to complement the quantitative findings.

8.2 Future Work

This dissertation advances the understanding of BWE by integrating empirical evidence, advanced analytical methods, and practical recommendations. The findings provide a structured approach to improving supply chain performance and resilience, emphasizing industry-specific strategies and innovative technologies to mitigate BWE's adverse effects. By contributing to both theoretical and practical domains, this research equips firms with tools to navigate modern supply chain complexities and optimize operational efficiency. Building on these insights, future research should explore several key areas.

Examining commonalities among industries presents an important research opportunity.

Understanding how industry characteristics influence BWE could lead to the development of predictive models that identify underlying patterns. Industries can be categorized into clusters based on production-distribution lead times, supply chain complexity, capacity scalability, and product durability. Analyzing the relationships between these clusters and supply chain performance can offer new perspectives on BWE mitigation strategies.

Further studies should explore the specific mechanisms through which importation affects supply chain variability and contributes to BWE. Identifying effective strategies to mitigate these impacts remains an essential research avenue. In particular, future research should assess the effectiveness of supply localization as a countermeasure to BWE by distinguishing between two critical dimensions: "persistent BWE" that is inherently tied to supply chain structures and "crisis-induced BWE" triggered by global disruptions such as pandemics. Empirical analyses utilizing panel data can help identify trends in BWE prevalence and its responsiveness to external shocks, offering a more nuanced understanding of supply chain behavior under different conditions.

The disruptions caused by the COVID-19 pandemic have highlighted the urgent need for agile and adaptive supply chain strategies. The lessons learned from this period should inform future research on long-term shifts in supply chain dynamics. The continued integration of advanced forecasting models and industry-level analyses could improve supply chain responsiveness to unforeseen disruptions. Investigating the sustained effects of the pandemic on supply chain operations and identifying strategies to enhance resilience will contribute to the development of robust, future-proof supply chains.

Future research should also focus on the scalability and real-world application of advanced forecasting techniques in mitigating BWE. The models proposed in this dissertation should be tested across different supply chain configurations and industries to evaluate their effectiveness in diverse operational settings. Integrating these forecasting techniques with decision-support systems and real-time data analytics could further enhance their ability to minimize demand amplification and improve overall supply chain efficiency.

Finally, expanding research on the relationship between supply chain collaboration and BWE mitigation could provide valuable insights. While this dissertation highlights the importance of information sharing, further exploration of collaboration mechanisms, such as digital supply chain platforms, blockchain-based transparency solutions, and AI-driven decision-making, could reveal new ways to dampen BWE and improve supply chain coordination.

9. REFERENCES

Abedsoltan, H. (2023) 'COVID-19 and the chemical industry: impacts, challenges, and opportunities', *Journal of Chemical Technology and Biotechnology*, 98(12). doi:10.1002/jctb.7531.

Agrawal, S., Sengupta, R.N. and Shanker, K. (2009) 'Impact of information sharing and lead time on bullwhip effect and on-hand inventory', *European Journal of Operational Research* [Preprint]. doi:10.1016/j.ejor.2007.09.015.

Anderson, E.G., Fine, C.H. and Parker, G.G. (2000) 'Upstream volatility in the supply chain: The machine tool industry as a case study', *Production and Operations Management*, 9(3), pp. 239–261. doi:10.1111/j.1937-5956.2000.tb00136.x.

Asadi, M. *et al.* (2023) 'Understanding interconnections among steel, coal, iron ore, and financial assets in the US and China using an advanced methodology', *International Review of Financial Analysis*, 89. doi:10.1016/j.irfa.2023.102789.

Avetisyan, M. *et al.* (2015) 'Competitiveness and macroeconomic impacts of reduced wait times at U.S. land freight border crossings', *Transportation Research Part A: Policy and Practice*, 78. doi:10.1016/j.tra.2015.04.027.

Baganha, M.P. and Cohen, M.A. (1998) 'The stabilizing effect of inventory in supply chains', *Operations Research*, 46(3-supplement-3), pp. S72–S83.

Baron, O., Callen, J.L. and Segal, D. (2018) 'Does the Bullwhip Matter Economically? A Cross-Sectional Firm-Level Analysis', *SSRN Electronic Journal* [Preprint], (972). doi:10.2139/ssrn.2856510.

Basu, R. and Ferreira, J. (2021) 'Sustainable mobility in auto-dominated Metro Boston:

Challenges and opportunities post-COVID-19', *Transport Policy*, 103. doi:10.1016/j.tranpol.2021.01.006.

Beason, R. (1993) 'Tests of production smoothing in selected Japanese industries', *Journal of Monetary Economics*, 31(3), pp. 381–394. doi:10.1016/0304-3932(93)90054-J.

Belhadi, A. *et al.* (2021) 'Manufacturing and service supply chain resilience to the COVID-19 outbreak: Lessons learned from the automobile and airline industries', *Technological Forecasting and Social Change*, 163. doi:10.1016/j.techfore.2020.120447.

Benrqya, Y. and Jabbouri, I. (2022) 'Impact of cross-docking on the bullwhip effect', *Journal of Modelling in Management*, ahead-of-print(ahead-of-print). doi:10.1108/JM2-03-2022-0088.

Bhattacharya, R. and Bandyopadhyay, S. (2011) 'A review of the causes of bullwhip effect in a supply chain', *International Journal of Advanced Manufacturing Technology* [Preprint]. doi:10.1007/s00170-010-2987-6.

Blanchard, O.J. (1983) 'The Production and Inventory Behavior of the American Automobile Industry', *Journal of Political Economy*, 91(3), pp. 365–400. doi:10.1086/261154.

Blinder, A.S. (1982) 'Inventories and Sticky Prices: More on the Microfoundations of Macroeconomics', *The American Economic Review*, 72(3), pp. 334–348.

Blinder, A.S. (1986) 'Can the production smoothing model of inventory behavior be saved?', *Quarterly Journal of Economics*, 101(3), pp. 431–453. doi:10.2307/1885691.

Blinder, A.S., Lovell, M.C. and Summers, L.H. (1981) 'Retail Inventory Behavior and Business Fluctuations', *Brookings Papers on Economic Activity*, 1981(2), pp. 443–520.

Blinder, A.S. and Maccini, L.J. (1991) 'Taking Stock: A Critical Assessment of Recent Research on Inventories', *Journal of Economic Perspectives*, 5(1), pp. 73–96. doi:10.1257/jep.5.1.73.

Bolarín, F.C., Frutos, A.G. and Mcdonnell, L.R. (2009) 'The Influence of Lead Time Variability on Supply Chain Costs: Analysis of Its Impact on the Bullwhip Effect.', *IUP*

Journal of Supply Chain Management, 6(3/4).

Bray, R.L. and Mendelson, H. (2012) ‘Information transmission and the bullwhip effect: An empirical investigation’, *Management Science*, 58(5), pp. 860–875. doi:10.1287/mnsc.1110.1467.

Bray, R.L. and Mendelson, H. (2015) ‘Production smoothing and the bullwhip effect’, *Manufacturing and Service Operations Management*, 17(2), pp. 208–220. doi:10.1287/msom.2014.0513.

Brinca, P., Duarte, J.B. and Faria-e-Castro, M. (2021) ‘Measuring labor supply and demand shocks during COVID-19’, *European Economic Review*, 139. doi:10.1016/j.euroecorev.2021.103901.

Buchmeister, B., Friscic, D. and Palcic, I. (2014) ‘Bullwhip Effect Study in a Constrained Supply Chain’, *Procedia Engineering*, 69, pp. 63–71. doi:10.1016/j.proeng.2014.02.204.

Bureau of Economic Analysis (2006) *What are intermediate inputs?*, *Bureau of Economic Analysis*.

Bureau of Transportation Statistics (2023) *Freight Analysis Framework*, *U.S. Department Of Transportation*. Available at: https://www.bts.gov/browse-statistical-products-and-data/state-transportation-statistics/transportation-economics (Accessed: 17 October 2023).

Cachon, G.P. (1999) ‘Managing supply chain demand variability with scheduled ordering policies’, *Management Science*, 45(6). doi:10.1287/mnsc.45.6.843.

Cachon, G.P., Randall, T. and Schmidt, G.M. (2007) ‘In search of the bullwhip effect’, *Manufacturing and Service Operations Management*, 9(4), pp. 457–479. doi:10.1287/msom.1060.0149.

Caplin, A.S. (1985) ‘The Variability of Aggregate Demand with (S, s) Inventory Policies’, *Econometrica*, 53(6), p. 1395. doi:10.2307/1913214.

Chaharsooghi, S.K. and Heydari, J. (2010) ‘LT variance or LT mean reduction in supply chain management: Which one has a higher impact on SC performance?’, *International Journal of Production Economics*, 124(2). doi:10.1016/j.ijpe.2009.12.010.

Chang, J.C. (1983) ‘An econometric model of the short-run demand for workers and hours in the U.S. auto industry’, *Journal of Econometrics*, 22(3), pp. 301–316. doi:10.1016/0304-4076(83)90106-9.

Chang, W.S. and Lin, Y.T. (2019) ‘The effect of lead-time on supply chain resilience performance’, *Asia Pacific Management Review*, 24(4). doi:10.1016/j.apmrv.2018.10.004.

Chatfield, D.C. *et al.* (2004) ‘The Bullwhip Effect-Impact of Stochastic Lead Time, Information Quality, and Information Sharing: A Simulation Study’, *Production and Operations Management*, 13(4), pp. 340–353. doi:10.1111/j.1937-5956.2004.tb00222.x.

Chen, F., Drezner, Z., *et al.* (2000) ‘Quantifying the bullwhip effect in a simple supply chain: the impact of forecasting, lead times, and information’, *Management Science*, 46(3), pp. 436–443. doi:10.1287/mnsc.46.3.436.12069.

Chen, F., Ryan, J.K. and Simchi-Levi, D. (2000) ‘The Impact of Exponential Smoothing Forecasts on the Bullwhip Effect’, *Naval Research Logistics*, 47(4). doi:10.1002/(SICI)1520-6750(200006)47:4<269::AID-NAV1>3.0.CO;2-Q.

Chen, L. and Lee, H.L. (2009) ‘Information sharing and order variability control Under a generalized demand model’, *Management Science*, 55(5). doi:10.1287/mnsc.1080.0983.

Chen, L. and Lee, H.L. (2012) ‘Bullwhip effect measurement and its implications’, *Operations Research*, 60(4), pp. 771–784. doi:10.1287/opre.1120.1074.

Chen, L. and Lee, H.L. (2017) ‘Modeling and Measuring the Bullwhip Effect’, in *Springer Series in Supply Chain Management*. doi:10.1007/978-3-319-32441-8_1.

Chen, L., Luo, W. and Shang, K. (2017) ‘Measuring the bullwhip effect: Discrepancy and alignment between information and material flows’, *Manufacturing and Service Operations Management*, 19(1). doi:10.1287/msom.2016.0590.

Cho, B. and Saki, Z. (2022) ‘Firm performance under the COVID-19 pandemic: The case of the U.S. textile and apparel industry’, *Journal of the Textile Institute*, 113(8). doi:10.1080/00405000.2021.1943258.

Cochrane, J.H. (1991) ‘A critique of the application of unit root tests’, *Journal of*

Economic Dynamics and Control, 15(2). doi:10.1016/0165-1889(91)90013-Q.

Cox, J.F., Blackstone, J.H. and Spencer, M.S. (2004) ‘American Production and Inventory Control Society (APICS) Dictionary’, *APICS Educational Society for Resource Manage* [Preprint].

Disney, S.M. *et al.* (2016) ‘Inventory management for stochastic lead times with order crossovers’, *European Journal of Operational Research*, 248(2). doi:10.1016/j.ejor.2015.07.047.

Djankov, S., Freund, C. and Pham, C.S. (2010) ‘Trading on time’, *Review of Economics and Statistics*, 92(1). doi:10.1162/rest.2009.11498.

Duc, T.T.H., Luong, H.T. and Kim, Y.D. (2008) ‘A measure of the bullwhip effect in supply chains with stochastic lead time’, *International Journal of Advanced Manufacturing Technology*, 38(11–12). doi:10.1007/s00170-007-1170-1.

Evans, C.L. and Harrigan, J. (2005) ‘Distance, time, and specialization: Lean retailing in general equilibrium’, *American Economic Review*. doi:10.1257/0002828053828590.

Fang, X. *et al.* (2013) ‘Decision support for lead time and demand variability reduction’, *Omega (United Kingdom)*, 41(2). doi:10.1016/j.omega.2012.03.005.

Fioriolli, J.C. and Fogliatto, F.S. (2008) ‘A Model to quantify the Bullwhip Effect in systems with stochastic demand and lead time’, in *2008 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2008*. doi:10.1109/IEEM.2008.4738040.

Forrester, J.W. (1961) *Industrial Dynamics*. MIT Press, Cambridge, Massachusetts.

Fransoo, J.C. and Wouters, M.J.F. (2000) ‘Measuring the bullwhip effect in the supply chain’, *Supply Chain Management*, 5(2), pp. 78–89. doi:10.1108/13598540010319993.

Gaalman, G., Disney, S.M. and Wang, X. (2022) ‘When bullwhip increases in the lead time: An eigenvalue analysis of ARMA demand’, *International Journal of Production Economics*, 250, p. 108623. doi:https://doi.org/10.1016/j.ijpe.2022.108623.

Ghansah, F.A. and Lu, W. (2023) ‘Responses to the COVID-19 pandemic in the construction industry: a literature review of academic research’, *Construction*

Management and Economics, 41(9). doi:10.1080/01446193.2023.2205159.

Haltiwanger, J.C. and Maccini, L.J. (1988) *A Model of Inventory and Layoff Behaviour Under Uncertainty*, *Source: The Economic Journal*.

Hamermesh Daniel S (1993) *Labor Demand*.

Hammond, J. (1994) *Barilla SpA*, *Harvard Business School Cases*. doi:10.2307/1353274.

Handfield, R.B., Graham, G. and Burns, L. (2020) 'Corona virus, tariffs, trade wars and supply chain evolutionary design', *International Journal of Operations and Production Management*, 40(10). doi:10.1108/IJOPM-03-2020-0171.

Herold, D.M. *et al.* (2021) 'COVID-19 and the pursuit of supply chain resilience: reactions and "lessons learned" from logistics service providers (LSPs)', *Supply Chain Management*, 26(6). doi:10.1108/SCM-09-2020-0439.

Heydari, J., Baradaran Kazemzadeh, R. and Chaharsooghi, S.K. (2009) 'A study of lead time variation impact on supply chain performance', *International Journal of Advanced Manufacturing Technology*, 40(11–12), pp. 1206–1215. doi:10.1007/s00170-008-1428-2.

Hochreiter, S. and Schmidhuber, J. (1997) 'Long Short Term Memory', *Neural Computation*, 9(8).

Holland, W. and Sodhi, M.S. (2004) 'Quantifying the effect of batch size and order errors on the bullwhip effect using simulation', *International Journal of Logistics*, 7(3), pp. 251–261. doi:10.1080/13675560412331298518.

Holt, C.C. *et al.* (1960) *Planning Production, Inventories, and Work Force*. Prentice-Hall (International series in management).

Holt, C.C., Modigliani, F. and Shelton, J.P. (1968) 'The Transmission of Demand Fluctuations Through a Distribution and Production System, the Tv-Set Industry', *The Canadian Journal of Economics*, 1(4), p. 718. doi:10.2307/133702.

Holt, C.C., Modigliani, F. and Simon, H.A. (1955) 'A Linear Decision Rule for Production and Employment Scheduling', *Management Science*, 2(1). doi:10.1287/mnsc.2.1.1.

Hummels, D.L. and Schaur, G. (2013) 'Time as a trade barrier', *American Economic Review*, 103(7). doi:10.1257/aer.103.7.2935.

International Monetary Fund (2009) *World Economic Outlook October 2009*.

Isaksson, O.H.D. and Seifert, R.W. (2016) 'Int . J . Production Economics Quantifying the bullwhip effect using two-echelon data : A cross-industry empirical investigation', *Intern. Journal of Production Economics*, 171, pp. 311–320. doi:10.1016/j.ijpe.2015.08.027.

Johnson, G.D. and Thompson, H.E. (1975) 'Optimality of Myopic Inventory Policies for Certain Dependent Demand Processes', *Management Science*, 21(11). doi:10.1287/mnsc.21.11.1303.

Kahn, J.A. (1987) 'Inventories and the Volatility of Production', *The American Economic Review*, 77(4), pp. 667–679. Available at: http://www.jstor.org/stable/1814538.

Kembro, J., Näslund, D. and Olhager, J. (2017) 'Information sharing across multiple supply chain tiers: A Delphi study on antecedents', *International Journal of Production Economics*, 193. doi:10.1016/j.ijpe.2017.06.032.

Kim, J.G. *et al.* (2006) 'Quantifying the bullwhip effect in a supply chain with stochastic lead time', *European Journal of Operational Research*, 173(2). doi:10.1016/j.ejor.2005.01.043.

Kowalski, C.J. (1972) 'On the Effects of Non-Normality on the Distribution of the Sample Product-Moment Correlation Coefficient', *Applied Statistics*, 21(1). doi:10.2307/2346598.

Lai, R.K. (2005) *Bullwhip in a Spanish Shop*, *Harvard NOM Working Paper*. doi:10.2139/ssrn.804745.

Lary, H.B. (1968) 'Trade in labor-intensive manufactures', in *Imports of manufactures from less developed countries*. NBER, pp. 86–115.

Lee, H.L., Padmanabhan, V. and Whang, S. (1997) 'Information Distortion in a Supply Chain: The Bullwhip Effect', *Management Science*, 43(4), pp. 546–558. doi:10.1287/mnsc.43.4.546.

Lee, H.L., So, K.C. and Tang, C.S. (2000) ‘The Value of Information Sharing in a Two-Level Supply Chain’, *Management Science*, 46(5), pp. 626–643. doi:10.1287/mnsc.46.5.626.12047.

Lee, S., Park, S.J. and Seshadri, S. (2022) ‘Variations of the Bullwhip Effect Across Foreign Subsidiaries’, *Manufacturing & Service Operations Management*, 25(1), pp. 1–18. doi:10.1287/msom.2022.1137.

Liu, Y. and Rabinowitz, A.N. (2021) ‘The impact of the COVID-19 pandemic on retail dairy prices’, *Agribusiness*, 37(1). doi:10.1002/agr.21687.

Mackelprang, A.W. and Malhotra, M.K. (2015) ‘The impact of bullwhip on supply chains: Performance pathways, control mechanisms, and managerial levers’, *Journal of Operations Management*, 36, pp. 15–32. doi:10.1016/j.jom.2015.02.003.

Maria del Rio-Chanona, R. *et al.* (2020) ‘Supply and demand shocks in the COVID-19 pandemic: An industry and occupation perspective’, *Oxford Review of Economic Policy*. doi:10.1093/oxrep/graa033.

Maritime Administration (2023) *U.S. Waterborne Foreign Trade by U.S. Customs Ports 2000-2017*, *U.S. Department Of Transportation*. Available at: https://www.maritime.dot.gov/data-reports/data-statistics/data-statistics (Accessed: 17 October 2023).

McCullen, P. and Towill, D. (2002) ‘Diagnosis and reduction of bullwhip in supply chains’, *Supply Chain Management*, 7(3), pp. 164–179. doi:10.1108/13598540210436612.

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

Michna, Z., Disney, S.M. and Nielsen, P. (2020) ‘The impact of stochastic lead times on the bullwhip effect under correlated demand and moving average forecasts’, *Omega (United Kingdom)*, 93. doi:10.1016/j.omega.2019.02.002.

Mollick, A.V. (2004) ‘Production smoothing in the Japanese vehicle industry’, *International Journal of Production Economics*, 91(1), pp. 63–74. doi:10.1016/j.ijpe.2003.07.001.

Nickell, S. (1984) ‘An Investigation of the Determinants of Manufacturing Employment in the United Kingdom’, *The Review of Economic Studies*, 51(4), pp. 529–557. doi:10.2307/2297778.

Nikolopoulos, K. *et al.* (2021) ‘Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions’, *European Journal of Operational Research*, 290(1). doi:10.1016/j.ejor.2020.08.001.

Parker, M.E.G. *et al.* (2021) ‘Public transit use in the United States in the era of COVID-19: Transit riders’ travel behavior in the COVID-19 impact and recovery period’, *Transport Policy*, 111. doi:10.1016/j.tranpol.2021.07.005.

Paul, S.K. and Chowdhury, P. (2020) ‘Strategies for Managing the Impacts of Disruptions During COVID-19: an Example of Toilet Paper’, *Global Journal of Flexible Systems Management*, 21(3). doi:10.1007/s40171-020-00248-4.

Pohlen, T.L. and Goldsby, T.J. (2003) ‘VMI and SMI programs: How economic value added can help sell the change’, *International Journal of Physical Distribution & Logistics Management*, 33(7), pp. 565–581.

Potter, A. and Disney, S.M. (2006) ‘Bullwhip and batching: An exploration’, *International Journal of Production Economics*, 104(2). doi:10.1016/j.ijpe.2004.10.018.

Ramsey, A.F. *et al.* (2021) ‘Impacts of COVID-19 and Price Transmission in U.S. Meat Markets’, *Agricultural Economics (United Kingdom)*, 52(3). doi:10.1111/agec.12628.

Rong, Y., Shen, Z.J.M. and Snyder, L. V. (2008) ‘The impact of ordering behavior on order-quantity variability: A study of forward and reverse bullwhip effects’, *Flexible Services and Manufacturing Journal*, 20(1–2), pp. 95–124. doi:10.1007/s10696-009-9054-3.

Rong, Y., Snyder, L. V. and Shen, Z.-J.M. (2008) ‘Bullwhip and Reverse Bullwhip Effects under the Rationing Game’, *SSRN Electronic Journal* [Preprint]. doi:10.2139/ssrn.1240173.

Rumyantsev, S. and Netessine, S. (2007) ‘What can be learned from classical inventory models? A cross-industry exploratory investigation’, *Manufacturing and Service Operations Management*, 9(4). doi:10.1287/msom.1070.0166.

Schmidt-Eisenlohr, T. (2013) ‘Towards a theory of trade finance’, *Journal of International Economics*, 91(1). doi:10.1016/j.jinteco.2013.04.005.

Seidu, R.K. *et al.* (2023) ‘Review of effects of COVID-19 pandemic on the textile industry: challenges, material innovation and performance’, *Research Journal of Textile and Apparel* [Preprint]. doi:10.1108/RJTA-08-2022-0098.

Shan, J. *et al.* (2014) ‘An empirical study of the bullwhip effect in China’, *Production and Operations Management*, 23(4), pp. 537–551. doi:10.1111/poms.12034.

Simon, H.A. (1952) ‘On the Application of Servomechanism Theory in the Study of Production Control’, *Econometrica*, 20(2), pp. 247–268. doi:10.2307/1907849.

Sterman, J.D. (1989) ‘Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment’, *Management Science*, 35(3), pp. 321–339. doi:10.1287/mnsc.35.3.321.

Suits, D.B. and Mack, R.P. (1958) ‘Consumption and Business Fluctuations. A Case Study of the Shoe, Leather, Hide Sequence.’, *Journal of the American Statistical Association*, 53(282). doi:10.2307/2281890.

Taylor, J.C., Robideaux, D.R. and Jackson, G.C. (2004) ‘U.S.-Canada transportation and logistics: Border impacts and costs, causes, and possible solutions’, *Transportation Journal*, 43(4).

The World Bank (2022) *United States Intermediate goods Imports By Country and Region*, *World Integrated Trade Solution (WITS)*. Available at: https://wits.worldbank.org/CountryProfile/en/Country/USA/StartYear/2007/EndYear/2007/TradeFlow/Import/Indicator/MPRT-TRD-VL/Partner/ALL/Product/UNCTAD-SoP2 (Accessed: 17 July 2022).

Verhoef, P.C., Noordhoff, C.S. and Sloot, L. (2023) ‘Reflections and predictions on effects of COVID-19 pandemic on retailing’, *Journal of Service Management*, 34(2). doi:10.1108/JOSM-09-2021-0343.

Walmsley, T. *et al.* (2023) ‘Macroeconomic consequences of the COVID-19 pandemic’, *Economic Modelling*, 120. doi:10.1016/j.econmod.2022.106147.

Wood, A. (1993) ‘Efficient Consumer Response’, *Logistics Information Management*, 6(4), pp. 38–40. doi:10.1108/EUM0000000002908.

Xiaojing, W. *et al.* (2008) ‘The impact of lead-time on bullwhip effect in supply chain’, *Proceedings - ISECS International Colloquium on Computing, Communication, Control, and Management, CCCM 2008*, 3, pp. 93–97. doi:10.1109/CCCM.2008.224.

Yao, Y., Duan, Y. and Huo, J. (2021) ‘On empirically estimating bullwhip effects: Measurement, aggregation, and impact’, *Journal of Operations Management*, 67(1), pp. 5–30. doi:10.1002/joom.1090.

Zhang, G.P. and Qi, M. (2005) ‘Neural network forecasting for seasonal and trend time series’, in *European Journal of Operational Research*. doi:10.1016/j.ejor.2003.08.037.

Zhao, R. *et al.* (2019) ‘Supply chain relational capital and the bullwhip effect: An empirical analysis using financial disclosures’, *International Journal of Operations and Production Management*, 39(5), pp. 658–689. doi:10.1108/IJOPM-03-2018-0186.

BIOGRAPHICAL SKETCH

Alper Sarıcıoğlu graduated from Işıklar Military High School, where he focused on math and science, and then earned a B.S. degree in Industrial Engineering from Eskişehir Osmangazi University in 2010. He later obtained his M.Sc. degree in Industrial Engineering from Bahçeşehir University in 2014 and is currently pursuing his Ph.D. in Industrial Engineering at Galatasaray University.

Alper has led numerous significant projects in logistics, supply chain optimization, warehousing strategy, and operational efficiency. His expertise includes network design, inventory management, and demand forecasting. Since September 2012, he has been a partner and senior consultant at Poseidon Consulting in Istanbul, where he has contributed to designing operations and developing IT strategies for various local and international companies. His responsibilities include project management, system analysis, and developing optimization applications for supply chain operations. His project portfolio also extends to engagements with Groupe PSA, Decathlon, SOCAR, Starbucks, Arzum, LCW, and several other prominent clients.

In addition to his consultancy work, Alper has contributed to academic research in supply chain management with his published study Analyzing One-Step and Multi-Step Forecasting to Mitigate the Bullwhip Effect and Improve Supply Chain Performance and his accepted study Impact of COVID-19 on the Bullwhip Effect Across U.S. Industries.