# Stores in an Omnichannel World: Understanding Their Role and Improving Their Performance

by

**Ayşe Çetinel**

A Dissertation Submitted to the
Graduate School of Business
in Partial Fulfillment of the Requirements for
the Degree of
Doctor of Philosophy
in

Operations Management and Information Systems

KOÇ ÜNİVERSİTESİ

July 17, 2024

**Stores in an Omnichannel World:**
**Understanding Their Role and Improving Their Performance**

Koç University

Graduate School of Business

This is to certify that I have examined this copy of a doctoral dissertation by

**Ayşe Çetinel**

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

______________________________________

Prof. A. Gürhan Kök (Co-advisor)

______________________________________

Assoc. Prof. Robert P. Rooderkerk (Co-advisor)

______________________________________

Assoc. Prof. Umut Güler

______________________________________

Assist. Prof. Murat Demirci

______________________________________

Prof. Abdurrahman B. Aydemir

______________________________________

Assoc. Prof. M. Berk Ataman

______________________________________

Assist. Prof. Burak Gökgür

Date: ______________________________

*To my son, Nova Kuzey,*

*Born in the last year of my studies,*

*May you always pursue your aspirations with courage,*
*no matter the stage of life you find yourself in.*
*May you explore the vast seas of life*
*with the clarity and guidance of the North Star.*

# ABSTRACT

**Stores in an Omnichannel World:**
**Understanding Their Role and Improving Their Performance**
**Ayşe Çetinel**
**Doctor of Philosophy in**
**Operations Management and Information Systems**
**July 17, 2024**

Online-first retailers are increasingly establishing physical stores to better serve their existing and new customers along the customer journey and increase customer engagement. This strategic expansion is driven by several key benefits. First, physical stores offer customers the invaluable opportunity to touch, feel, and try products before buying, significantly enriching the customer experience and reducing hesitation in purchase decisions—especially crucial for products where sensory experiences influence buying behavior. Furthermore, offline stores can enhance brand perception or act as a billboard for the brand's existence, thus enhancing visibility and credibility in a competitive marketplace. Additionally, these stores improve customer service by providing immediate, face-to-face assistance, which can swiftly resolve issues and facilitate returns and exchanges, thereby boosting customer satisfaction and loyalty. By allowing customers to verify product quality and fit before purchase, physical stores can also substantially lower return rates, reducing the logistical and financial burdens associated with handling returns. Besides serving existing customers better, physical stores can also capture additional sales opportunities by attracting different customer segments, including those who prefer in-store shopping or who engage in impulse buying when physically present in a store. Lastly, the synergy between online and offline channels enhances the overall shopping experience, enabling features like buy online, pick up in-store, which provides consumers with convenient, flexible shopping options. This omnichannel approach not only meets the diverse preferences of modern consumers but also drives greater brand loyalty and market reach.

Building on these insights, this thesis explores the strategic benefits and challenges of stores in an omnichannel business model across three comprehensive chapters. The first chapter, "The Value of Experience-centric Stores in Omnichannel Retail," investigates the operational and financial implications of opening physical stores. It specifically analyzes three omnichannel store openings of an online-first reseller of consumer electronics - two large experience-centric stores and one small city-center format. Contrary to prior research, our results reveal that physical retail expansion, either by experience-centric or small stores, does not yield a significant positive effect toward online sales. Rather, we find patterns of cannibalization—quite significant for one of the two large stores. Revenues generated by the small store failed to offset sales lost in the

online channel. For the large experience-centric stores, however, we reveal increases in total net revenues in the range of 21% - 23% after opening, rising further long term.

Utilizing a diverse range of product categories and data from an online consumer survey, we explore category-specific insights on the value of experience-centric stores in omnichannel retail. The survey assesses the perceived utility of stores across three phases of the customer journey: information search, fulfillment, and product returns. Our results show that our utility-based framework effectively captures the short-term and, to a lesser degree, long-term value that stores offer to both retailers and consumers, particularly in destination categories. Store openings lead to greater overall revenue increases in product categories where consumers find higher utility in the physical store experience. However, the framework is less effective for accessory categories, suggesting a need for additional investigation in this area. These findings provide retailers with crucial insights into which product categories benefit most from experience-centric stores and enhance the understanding of the interaction between online and offline channels in an omnichannel retail environment.

The second chapter, "The Effect of Omnichannel Store Openings on Return Rates," investigates how physical stores impact return behavior a significant concern for online-first retailers. As product returns escalate, online retailers have tightened their policies to curb losses, whereas their physical counterparts generally experience lower return rates due to customers' ability to inspect products and consult with staff. This contrast forms the backdrop for online-first retailers expanding into brick-and-mortar under an omnichannel model, aiming to reduce return rates by leveraging the tactile benefits of physical stores. However, this expansion could increase returns as physical locations might also increase the convenience of returning.

This chapter introduces a conceptual framework to examine how physical store openings affect net revenue uplift by analyzing the interplay between increased sales and returned revenues. The findings reveal that while physical stores can enhance customer experiences and satisfaction, they may inadvertently lead to higher item return rates, which rose by 28.0% for Large Store 1 and 21.7% for Large Store 2. This highlights that the effectiveness of omnichannel strategies relies not just on the presence of physical stores but on their seamless integration with online operations to influence consumer behavior and manage returns effectively. The analysis shows that the higher average price of returned items compared to sold items amplifies the effect of increased return rates, leading to a more significant uplift in the return fraction. Consequently, net revenue grows less than gross revenue, illustrating the delicate balance between increasing revenue and controlling return rates. Given the retailer's low return rate before the store openings, the store-level return rates and return fractions increased by only about 0.5 percentage points. This can be considered the 'cost of doing business,' yet it remains a crucial aspect to monitor and mitigate.

The third chapter, "Rightsizing Store Labor," focuses on ways to improve store productivity by evaluating adaptive staffing strategies in omnichannel retail. Given the labor-intensive nature of omnichannel stores—where activities such as click-and-collect, returns, in-store advice, purchases, and fulfillment are integral—efficient labor management is critical. This chapter particularly examines the potential for retailers to dynamically adjust labor hours in real-time based on actual customer traffic, with adjustments made voluntarily by employees. This analysis covers both scenarios where retailers can enhance labor productivity during low traffic periods by reducing hours without sacrificing sales, and where they can capitalize on unexpected customer traffic spikes by increasing hours to potentially boost sales. A field experiment was conducted to assess the impact of these labor adjustments on store productivity. We examined how the autonomy to scale hours impacts productivity and investigated the specifics of how the retailer utilizes this flexibility, the types of adjustments made, and the resulting effects on overall productivity.

The intent-to-treat results indicated a significant 6.24% increase in productivity, with negligible impacts on labor hours and sales. More strikingly, the average treatment effect on the treated revealed significant enhancements: a 17.6% increase in productivity and a 10.3% reduction in labor hours when staffing was scaled up. These findings highlight the importance of real-time staffing flexibility in correcting planning imperfections and boosting overall performance. Upscaling labor hours during peak periods led to significant productivity gains (*effectiveness* route), while downscaling during low traffic periods had an insignificant effect (*efficiency* route). The 10.3% reduction in labor hours when staffing was scaled up suggests more efficient labor use, illustrating how on-the-day scaling can be used to (partially) correct for imperfect labor allocations made well in advance when store traffic estimates are full of uncertainty. This indicates that stores were initially planned with fewer labor hours to match actual traffic, and despite the upscaling, they did not fully close the labor hour gap, implying more efficient labor utilization.

These insights are crucial for omnichannel retailers facing high inflation and rising labor costs. Dynamic labor management enables retailers to respond effectively to fluctuating customer traffic, ensuring optimal staffing levels that enhance customer satisfaction and operational efficiency. This study provides actionable strategies for maintaining operational efficiency, offering a vital approach for retailers to manage labor resources effectively amidst economic pressures. By integrating these dynamic staffing strategies, retailers can better balance operational efficiency and labor management, navigating today's retail landscape and improving the synergy between online and offline channels.

Together, these chapters provide a holistic view of the strategic, financial, and operational outcomes of integrating online and offline retail channels, offering crucial insights for retailers navigating the challenges of today's retail landscape.

# ÖZETÇE

**Çok Kanallı Dünyada Mağazaların**
**Rollerini Anlama ve Performanslarını İyileştirme Üzerine**
**Ayşe Çetinel**

**Operasyon Yönetimi ve Bilgi Sistemleri, Doktora**
**17 Temmuz, 2024**


Çevrimiçi öncelikli perakendeciler, müşteri yolculuğu boyunca mevcut ve yeni müşterilerine daha iyi hizmet vermek ve müşteri etkileşimini artırmak için giderek daha fazla fiziksel mağaza açıyor. Bu stratejik genişleme birçok önemli faydadan kaynaklanmaktadır. Birincisi, fiziksel mağazalar müşterilere satın almadan önce ürünlere dokunma, hissetme ve deneme konusunda paha biçilmez bir fırsat sunarak müşteri deneyimini önemli ölçüde zenginleştiriyor ve satın alma kararlarındaki tereddütleri azaltıyor; özellikle duyusal deneyimlerin satın alma davranışını etkilediği ürünler için çok önemli. Ayrıca, çevrimdışı mağazalar marka algısını artırabilir veya markanın varlığı için bir reklam panosu görevi görebilir, böylece rekabetçi bir pazarda görünürlüğü ve güvenilirliği artırabilir. Ayrıca bu mağazalar, sorunları hızlı bir şekilde çözebilen, iade ve değişimleri kolaylaştırabilen, anında yüz yüze yardım sağlayarak müşteri hizmetlerini geliştiriyor ve böylece müşteri memnuniyetini ve sadakatini artırıyor. Fiziksel mağazalar, müşterilerin satın almadan önce ürün kalitesini ve uygunluğunu doğrulamalarına olanak tanıyarak iade oranlarını da önemli ölçüde düşürebilir ve iade işlemleriyle ilgili lojistik ve mali yükleri azaltabilir. Fiziksel mağazalar, mevcut müşterilere daha iyi hizmet vermenin yanı sıra, mağaza içi alışverişi tercih edenler veya bir mağazada fiziksel olarak bulunduklarında anlık satın alma gerçekleştirenler de dahil olmak üzere farklı müşteri segmentlerini çekerek ek satış fırsatları yakalayabilir. Son olarak, çevrimiçi ve çevrimdışı kanallar arasındaki sinerji, tüketicilere kullanışlı, esnek alışveriş seçenekleri sunan çevrimiçi satın alma, mağazadan teslim alma gibi özellikleri etkinleştirerek genel alışveriş deneyimini geliştiriyor. Bu çok kanallı yaklaşım, yalnızca tüketicilerin farklı tercihlerini karşılamakla kalmıyor, aynı zamanda daha fazla marka bağlılığı ve pazar erişimi sağlıyor.

Bu bilgilerden yola çıkarak bu tez, çok kanallı bir iş modelinde mağazaların stratejik faydalarını ve zorluklarını üç kapsamlı bölümde araştırıyor. "Çok Kanallı Perakendede Deneyim Odaklı Mağazaların Değeri" başlıklı ilk bölüm, fiziksel mağaza açmanın operasyonel ve finansal sonuçlarını inceliyor. Tüketici elektroniği alanında ilk çevrimiçi satıcının üç çok kanallı mağaza açılışını özellikle analiz ediyor: iki büyük deneyim odaklı mağaza ve bir küçük şehir merkezi formatı. Önceki araştırmaların aksine sonuçlarımız, ister deneyim odaklı ister küçük mağazalar aracılığıyla olsun fiziksel perakende genişlemesinin, çevrimiçi satışlar üzerinde önemli bir olumlu etki yaratmadığını ortaya koyuyor. Aksine, iki büyük mağazadan biri için oldukça önemli olan yamyamlaştırma

(cannibalization) modellerini buluyoruz. Küçük mağazanın elde ettiği gelirler, çevrimiçi kanalda kaybedilen satışları telafi edemediğini; ancak deneyim odaklı büyük mağazalar için, açılıştan sonra toplam net gelirlerde %21 - %23 aralığında artışlar olduğunu ve uzun vadede daha da arttığını ortaya koyuyoruz.

Çok çeşitli ürün kategorilerinden ve çevrimiçi tüketici anketinden elde edilen verilerden yararlanarak, çok kanallı perakendede deneyim odaklı mağazaların değerine ilişkin kategoriye özel içgörüleri araştırıyoruz. Anket, müşteri yolculuğunun üç aşamasında mağazaların algılanan faydasını değerlendiriyor: bilgi arama, sipariş karşılama ve ürün iadeleri. Sonuçlarımız, fayda temelli çerçevemizin, özellikle de destinasyon kategorilerinde mağazaların hem perakendecilere hem de tüketicilere sunduğu kısa vadeli ve daha az ölçüde uzun vadeli değeri etkili bir şekilde yakaladığını gösteriyor. Mağaza açılışları, tüketicilerin fiziksel mağaza deneyiminden daha fazla faydalandığı ürün kategorilerinde daha fazla genel gelir artışına yol açıyor. Ancak fayda temelli çerçevenin aksesuar kategorileri için daha az etkili olması, bu alanda ilave araştırmalara ihtiyaç duyulduğunu ortaya koymaktadır. Bu bulgular perakendecilere, deneyim odaklı mağazalardan en çok hangi ürün kategorilerinin yararlandığına dair önemli bilgiler sağlıyor ve çok kanallı bir perakende ortamında çevrimiçi ve çevrimdışı kanallar arasındaki etkileşimin anlaşılmasını geliştiriyor.

"Çok Kanallı Mağaza Açılışlarının İade Oranları Üzerindeki Etkisi" başlıklı ikinci bölüm, fiziksel mağazaların, çevrimiçi öncelikli perakendeciler için önemli bir endişe olan iade davranışını nasıl etkilediğini araştırıyor. Ürün iadeleri arttıkça, çevrimiçi perakendeciler kayıpları azaltmak için politikalarını sıkılaştırırken, fiziki muadilleri, müşterileri ürünleri inceleyebildikleri ve personelden danışmanlık alabilecekleri için daha düşük iade oranlarına sahiptir. Bu tezat, çevrimiçi perakendecilerin çok kanallı bir model altında fiziksel mağazalara genişleyerek iade oranlarını azaltma çabalarının arka planını oluşturur, çünkü fiziksel mağazaların dokunsal faydalarından yararlanmayı hedeflerler. Ancak, bu genişleme iade oranlarını artırabilir çünkü fiziksel mağazalar iade işlemlerini de kolaylaştırabilmektedirler.

Bu bölüm, artan satışlar ile iade edilen gelirler arasındaki etkileşimi analiz ederek fiziksel mağaza açılışlarının net gelir artışını nasıl etkilediğini incelemek için kavramsal bir çerçeve sunmaktadır. Bulgular, fiziksel mağazaların müşteri deneyimlerini ve memnuniyetini artırabildiğini ancak istemeden daha yüksek ürün iade oranlarına yol açabileceğini ortaya koyuyor. Birinci büyük mağaza için %28,0 ve ikinci büyük mağaza için %21,7 artış göstermiştir. Bu, çok kanallı stratejilerin etkinliğinin yalnızca fiziksel mağazaların varlığına değil, aynı zamanda çevrimiçi operasyonlarla sorunsuz entegrasyonuna bağlı olduğunu vurgulamaktadır. Analiz, satılan ürünlere kıyasla iade edilen ürünlerin daha yüksek ortalama fiyatının, artan iade oranlarının etkisini artırdığını göstermektedir ve bu da iade oranındaki artışın daha büyük bir gelir kaybına yol açtığını göstermektedir. Sonuç olarak, net gelir, brüt gelirden daha az artmakta ve bu da gelir artışı ile iade oranlarını kontrol etme arasındaki hassas dengeyi göstermektedir. Perakendecinin mağaza açılışlarından önce düşük iade oranı göz önüne alındığında, mağaza düzeyinde iade oranları ve iade yüzdeleri sadece yaklaşık %0,5 puan artmıştır. Bu, "iş yapmanın maliyeti" olarak kabul edilebilir, ancak izlenmesi ve azaltılması gereken önemli bir yön olarak kalmaktadır.

"Mağaza İşgücünün Doğru Boyutlandırılması" başlıklı üçüncü bölüm, çok kanallı perakendede uyarlanabilir personel stratejilerini değerlendirerek mağaza verimliliğini artırmanın yollarına odaklanıyor. Tıkla ve topla, iadeler, mağaza içi tavsiyeler, satın almalar ve sipariş karşılama gibi faaliyetlerin ayrılmaz bir parçası olduğu çok kanallı mağazaların emek yoğun doğası göz önüne alındığında, verimli iş gücü yönetimi kritik öneme sahiptir. Bu bölüm, özellikle perakendecilerin, çalışanların gönüllü olarak yaptığı ayarlamalarla, gerçek müşteri trafiğine dayalı olarak çalışma saatlerini gerçek zamanlı olarak dinamik olarak ayarlama potansiyelini inceliyor. Bu analiz, perakendecilerin trafiğin düşük olduğu dönemlerde satışlardan ödün vermeden saatleri azaltarak işgücü verimliliğini artırabilecekleri ve satışları potansiyel olarak artırmak için saatleri artırarak beklenmedik müşteri trafiği artışlarından yararlanabilecekleri her iki senaryoyu da kapsar. Bu işgücü düzenlemelerinin mağaza verimliliği üzerindeki etkisini değerlendirmek için bir saha deneyi yapıldı. Saatleri ölçeklendirme özerkliğinin üretkenliği nasıl etkilediğini inceledik ve perakendecinin bu esnekliği nasıl kullandığının ayrıntılarını, yapılan ayarlama türlerini ve bunun genel üretkenlik üzerindeki etkilerini araştırdık.

Uygulama niyetine dayalı (intent-to-treat) sonuçlar, çalışma saatleri ve satışlar üzerinde ihmal edilebilir etkilerle birlikte %6,24'lük önemli bir verimlilik artışı gösterdi. Daha çarpıcı bir şekilde, tedavi edilenler üzerindeki ortalama tedavi etkisi (average treatment effect on the treated) önemli iyileşmeler ortaya çıkardı: personel sayısı artırıldığında verimlilikte %17,6 artış ve çalışma saatlerinde %10,3 azalma. Bu bulgular, planlama kusurlarının düzeltilmesinde ve genel performansın artırılmasında gerçek zamanlı personel esnekliğinin önemini vurgulamaktadır. Yoğun dönemlerde çalışma saatlerinin artırılması önemli üretkenlik artışlarına yol açarken (*etkinlik yolu*), trafiğin düşük olduğu dönemlerde ölçeğin düşürülmesi ise önemsiz bir etkiye (*verimlilik yolu)* yol açtı. Personel sayısı artırıldığında çalışma saatlerindeki %10,3'lük azalma, daha verimli işgücü kullanımına işaret ediyor; bu da, mağaza trafiği tahminlerinin belirsizliklerle dolu olduğu durumlarda çok önceden yapılan kusurlu işgücü tahsislerini (kısmen) düzeltmek için günlük ölçeklendirmenin nasıl kullanılabileceğini gösteriyor. Bu, mağazaların başlangıçta gerçek trafiğe uyacak şekilde daha az çalışma saati ile planlandığını ve ölçeklendirmeye rağmen çalışma saati açığını tam olarak kapatmadıklarını, bu da daha verimli iş gücü kullanımı anlamına geldiğini gösteriyor.

Bu bilgiler, yüksek enflasyon ve artan işgücü maliyetleriyle karşı karşıya kalan çok kanallı perakendeciler için hayati önem taşıyor. Dinamik iş gücü yönetimi, perakendecilerin dalgalanan müşteri trafiğine etkili bir şekilde yanıt vermesini sağlayarak müşteri memnuniyetini ve operasyonel verimliliği artıran optimum personel seviyelerini sağlar. Bu çalışma, operasyonel verimliliği sürdürmek için eyleme geçirilebilir stratejiler sunarak perakendecilerin ekonomik baskıların ortasında iş gücü kaynaklarını etkili bir şekilde yönetmeleri için hayati bir yaklaşım sunuyor. Perakendeciler, bu dinamik personel stratejilerini entegre ederek operasyonel verimlilik ile işgücü yönetimini daha iyi dengeleyebilir, günümüz perakende ortamında gezinebilir ve çevrimiçi ve çevrimdışı kanallar arasındaki sinerjiyi geliştirebilir.

Bu bölümler bir arada, çevrimiçi ve çevrimdışı perakende kanallarını entegre etmenin stratejik, finansal ve operasyonel sonuçlarına ilişkin bütünsel bir bakış sunarak günümüz perakende ortamının zorluklarıyla başa çıkan perakendecilere önemli bilgiler sunuyor.

# ACKNOWLEDGMENTS

I would like to express my deepest gratitude to everyone who supported me throughout my journey in completing this dissertation.

First and foremost, I am immensely grateful to my advisors, Prof. Gürhan Kök from Koç University and Assoc. Prof. Robert Rooderkerk from Rotterdam School of Management (RSM), Erasmus University. Gürhan, who conducted my graduate school interview process for admission to the GSB program and later became my advisor, was the main reason I chose to join Koç. His introduction to Robert when I expressed interest in empirical analysis with big data at the marketing-operations interface was pivotal. Gürhan has been a constant example of hard work, passion for research, and respect, making an inestimable contribution to my training as a scholar and as a person.

Robert, with his great sense of humor and ability to demonstrate work-life balance, has been a role model. He spent considerable time providing feedback on all my research, pushing our first paper to receive a minor revision in the first round at the Production and Operations Management journal. His talent for storytelling, ability to place research in a business context, and exemplary presentation skills have profoundly influenced my work. Robert's support extended beyond academics, consistently bringing opportunities for professional and personal growth to my attention. My three months working with him at RSM were incredibly enriching. He often reminded me that this journey is not a big marathon but a series of sprints. Robert's constant effort to make academic life fun and relevant has been invaluable. Moreover, he facilitated our industry partnership, which significantly enriched this dissertation.

Both Gürhan and Robert have been pivotal figures in my academic journey. Their guidance and unwavering support have been instrumental in shaping this research and my professional development.

I am deeply grateful to our industry partner, the online-first consumer electronics retailer in Western Europe, for their collaboration and patience while we focused on academic rigor, which inevitably takes time. Their willingness to share insights and experiences has greatly enriched this research.

I am also thankful to my committee members for their constructive feedback and suggestions: Assoc. Prof. Umut Güler, Assist. Prof. Murat Demirci, Prof. Abdurrahman Aydemir, Assoc. Prof. Berk Ataman, and Assist. Prof. Burak Gökgür. Special thanks to Assist. Prof Murat Demirci, who spent time with me for Chapter 3, offering his invaluable insights and support.

A special thanks to my friends, Deniz and Bilal, whom I met in econometrics courses at Koç University and who have become my life-long friends. They have always been there, providing much-needed emotional support and offering invaluable assistance, discussions, and feedback. A number of people made this long journey a little bit less lonely for me. These include my friends Elif, Alireza, and Onur in the doctoral program from Koç University, and Roger and Begüm from RSM.

When I began my PhD, I had already spent over a decade working in Silicon Valley. Initially, my parents had some reservations about my pursuit of an academic career. However, once I started, they were fully on board and provided me with unwavering support.

Last but certainly not least, I would like to dedicate this work to my family, who have waited patiently for me during endless days. My deepest appreciation goes to Efe for his unwavering love, patience, and encouragement. During my last year of the PhD, we welcomed Nova Kuzey into the world. Thank you, Efe, for taking such good care of us, and for making my absence less noticeable to our son. You have been my anchor, providing stability, and my compass, showing me the direction when waters get choppy. Nova Kuzey, now almost 10 months old, has brought immense joy and motivation into our lives. Together, we have made this possible.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# INTRODUCTION

In the retail landscape, omnichannel strategies have become increasingly vital as retailers strive to seamlessly integrate online and offline experiences. A typical shopping journey might begin online, transition to an in-store visit to inspect a product, and conclude with a purchase made on a mobile device. This seamless integration of shopping channels is the essence of the omnichannel journey, driven by the unique strengths of each channel and enhancing the overall customer experience.

To remain consistently customer-centric, retailers are integrating these channels, thereby improving our shopping experience. This is why even online retailers are opening physical stores—to create a seamless and convenient customer experience. Physical stores allow customers to inspect products in person and receive advice from store employees. Additionally, customers can pick up or return online purchases in-store. This complementarity between online and physical stores has motivated online retailers to open physical locations, resulting in a surge in net store openings over the last two years. The rapid expansion of physical stores by digitally native vertical brands (DNVBs) and online multi-brand retailers highlights a strategic shift aimed at enhancing customer experience and operational efficiency.

However, expanding into physical retail is not without its challenges. It requires a completely different set of skills, including leasing and running store locations, organizing supply chains and logistics for stores, handling store returns, and offering services like buy online, pick up in-store. This shift illustrates what goes into becoming an omnichannel retailer and highlights the need for strategic, tactical, and operational decisions.

Online retailers must first determine their optimal retail format, deciding which store format to open and where. It took many years for traditional retailers like Target to find a strategy that works. These retailers are not merely increasing their physical presence but are also strategically designing stores to address the limitations of the online customer journey, offering tangible product interactions, expert guidance, and efficient fulfillment and return processes (Avery et al., 2012; Kapner, 2021).

Tactical decisions, such as managing returns, are also critical. Opening stores can potentially reduce returns, as customers can physically inspect the products. However, it can also lead to increased returns since customers can easily return any purchase in-store, including online purchases. The challenge of reducing return rates while maintaining customer satisfaction illustrates the complexity of omnichannel retail management.

Operational challenges include managing store operations, which are costly and difficult. Payroll constitutes almost 10% of sales, making it crucial to focus on labor productivity (Kesavan & Mani, 2015). Otherwise, cutting jobs becomes inevitable. Effective labor management is also critical, as maintaining optimal staffing levels directly impacts customer service quality and operational efficiency. Balancing staffing costs with service levels is essential for converting customer traffic into sales and fostering long-term customer loyalty (Kesavan & Mani, 2015; Lee, 2023). The implementation of voluntary real-time staffing adjustments, similar to just-in-time scheduling, offers a promising solution by dynamically aligning labor supply with actual customer demand, thereby enhancing store productivity and employee satisfaction (Kamalahmadi et al., 2021).

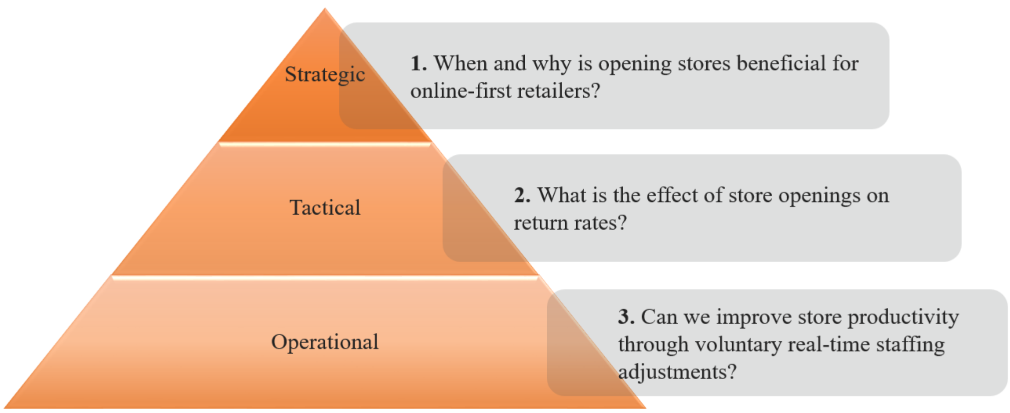
This body of research collectively explores the multifaceted role of physical stores for online-first retailers, divided into three key studies, each addressing a different dimension. In Chapter 1, we address a strategic question: "When and why do online multi-brand retailers benefit from opening physical stores?" This chapter provides empirical insights into how these stores offer unique in-store experiences that complement online shopping. By analyzing empirical data, this chapter extends the understanding of strategic decisions related to store openings and their impact on the overall retail strategy. This research was funded by the Marketing Science Institute and has received a minor revision in the first round at the Production and Operations Management journal.

Chapter 2 investigates a tactical question on returns: "What is the impact of physical store openings on return rates?" By analyzing empirical data, this chapter examines whether the introduction of physical stores can mitigate the high return rates associated with online shopping. The findings highlight the importance of effective return rate management to mitigate its impact on net revenue uplift.

Chapter 3 addresses the operational dimension: "Can store productivity be improved through voluntary intraday staffing adjustments?" This chapter presents a field experiment on real-time staffing adjustments, demonstrating how dynamically aligning labor supply with actual customer demand can enhance store productivity. The insights from this chapter highlight the importance of balancing staffing costs with service levels to optimize overall store performance.

For all these studies, we partnered with the same online-first consumer electronics retailer in Western Europe. Our retail partner, like Apple, opens large experience-centric stores, focusing on product showcasing and customer interactions. Currently, they operate 15 stores and aim to reach 50 by the end of next year. Their omnichannel strategies prioritize customer-centric, high-quality service. They tailor product journeys to customer needs, and their high Net Promoter Score reflects their success in customer satisfaction.

**Figure** Understanding stores' role and improving their performance through three studies




By exploring the value of experience-centric stores, the effect of store openings on return rates, and innovative labor management strategies, these studies offer valuable insights into optimizing omnichannel operations to drive growth and enhance customer experience. The subsequent chapters systematically investigate the integration of physical stores into an omnichannel retail strategy, offering a holistic understanding of its challenges and opportunities. The findings contribute to academic knowledge and offer practical insights for retailers seeking to drive growth and enhance customer experience in a rapidly evolving retail landscape. This dissertation emphasizes the ongoing need for innovation and adaptability, highlighting the importance of a holistic approach to retail management that seamlessly integrates online and offline experiences.

Chapter 2

# THE EFFECT OF OMNICHANNEL STORE OPENINGS ON RETURN RATES

## 2.1 Introduction

The pandemic has dramatically reshaped the retail landscape, compelling retailers to implement lenient return policies to promote online shopping. As consumers shifted to online purchases, return rates soared, leading to unprecedented challenges for retailers. The surge in online returns results in significant costs in handling and loss of product value (Phillips, 2017; Ofek et al., 2010). In response, several online retailers have tightened their return policies to improve profitability and operational sustainability (Kapner, 2023). For example, Zara has responded by shortening refund windows and introducing restocking fees (Wolfe, 2022). According to the National Retail Federation, the online return rate has escalated from 11% to 18%, while in-store returns remain significantly lower due to the tactile experience and personal advice available to customers before making a purchase (NRF, 2023).

This trend presents a contradiction for online retailers considering physical store expansions. On the one hand, online retailers might open physical stores expecting to reduce return rates by offering hands-on experiences and personalized assistance, enhancing customer satisfaction and potentially lowering return rates. However, physical stores might inadvertently facilitate higher returns by making the process more convenient, such as accepting online returns in-store at no extra cost. Notably, nearly half of in-store returns originate from online purchases (NRF, 2023). This dichotomy guides our research to investigate whether expanding into physical stores can effectively mitigate return rates for an online-first retailer.

Research on the impact of physical stores on returns is limited, yet critical. We specifically explore how such expansions influence different metrics of return rates, including the *item return rate* (referred as *return rate*) and *revenue return rate* (referred as *return fraction*). We propose a conceptual framework to illustrate how changes in the return fraction—the fraction of revenues that are returned—affect the net revenue uplift and how this, in turn, is influenced by changes in the return rate, average price per unit sold, and average price per item returned.

Our research setting involves an online-first consumer electronics retailer opening physical stores across a diverse range of product categories with varying return rates. Physical stores can critically influence return rates through in-store customer service and hands-on product experiences. This scenario offers a valuable opportunity to study the potential of physical stores to reduce return rates for an online-first retailer.

Our findings indicate that the introduction of physical stores has impacted returns, with a noticeable shift in consumer behavior towards more frequent returns. We observe a significant increase in return revenues stemming from an overall rise in return rates post-store opening, suggesting that stores may be making the return process more convenient, thereby impacting the retailer's operational strategy. Specifically, return rates rose by 28% for Large Store 1 and 21.7% for Large Store 2. The analysis further reveals that the higher average price of returned items compared to sold items amplifies the effect of increased return rates, resulting in a more pronounced uplift in the return fraction. Consequently, net revenue grows less than gross revenue, highlighting the delicate balance between increasing revenue and controlling return rates.

## 2.2 Literature review

### 2.2.1 Introduction

*Overview of item return rates and revenue return rates.* Return rates can be categorized into *item return rates* (referred to as *return rate*) and *revenue return rates* (referred to as *return fraction*). Item return rates, referring to the percentage of returned items after purchase, impact inventory management, operational costs, and customer satisfaction. High item return rates lead to lost sales revenue and additional costs for restocking and processing returns, thereby affecting the retailer's bottom line (Lee & Morewedge, 2023).

Revenue return rates focus on the financial impact of returns, including lost sales revenue, restocking costs, and overall profitability.

*Comparison of return rates.* The disparity between these two return rates arises from differences in the average price of returned items versus purchased items. They align when the average price of a returned item is the same as the average price of a purchased item. However, if more expensive products have a higher return rate, the revenue return rate will exceed the item return rate. Aligning return policies with consumer expectations and implementing effective strategies to reduce return rates enhances customer satisfaction and maximizes revenue (Kong et al., 2017; Yoon & Jeong, 2021).

*Online versus physical stores.* In the retail sector, the disparities in return rates between online and physical stores are well-documented and represent a significant operational challenge. Li and Liu (2021) report that return rates for online retailers are typically double those of brick-and-mortar stores. This is supported by Xia et al. (2016), who found that online purchases have return rates ranging from 8% to 11%, significantly higher than the less than 9% observed in physical stores. These high return rates are particularly impactful financially, as Lee and Morewedge (2023) illustrate with data showing that in 2022, returns amounted to $816 billion in lost revenue for retailers, marking a significant increase from previous years.

### 2.2.2 Financial and operational impact

The operational and financial repercussions of these high online return rates are critical for retailers. Bernon et al. (2016) emphasize that managing these returns efficiently is critical, as the cost and challenges of handling online returns are considerably higher than for physical stores. This situation is complicated by inventory and supply chain challenges, as high return rates can lead to significant backlogs, affecting both inventory levels and operational efficiency (Liu & Xu, 2020). Furthermore, the nature of online shopping contributes to higher return rates due to factors such as the inability to physically inspect products before purchase, leading to mismatches in customer expectations (Shang et al., 2017). This highlights the importance of developing strategic approaches to minimize returns, such as enhancing product descriptions, improving sizing information, and using high-quality visuals.

### 2.2.3 Role of physical stores

Integrating physical stores into the retail mix can significantly reshape the patterns of returns. Gao and Su (2017) illustrate that physical stores can serve as showrooms, decreasing the likelihood of returns by allowing customers to inspect and evaluate products before purchasing. This tactile interaction helps reduce the uncertainty often associated with online shopping. Similarly, Akturk et al. (2018) note that introducing ship-to-store services can merge the convenience of online shopping with the assurance provided by physical stores, increasing sales and decreasing returns. The potential of physical stores to reduce return rates is corroborated by Hirche et al. (2022), who observed that direct interactions in physical stores provide immediate customer service and tactile experiences, thereby reducing overall return rates. This presents a strategic advantage as physical stores can address some inherent challenges of online retailing by enhancing consumer confidence at the point of purchase. Li (2023) adds that the personalized service in physical stores, where salespeople provide current and relevant information rather than overwhelming consumers with excessive details, can significantly improve the shopping experience. Such tailored interactions can lead to more informed purchasing decisions, thus reducing the propensity for returns. These insights collectively highlight the transformative impact that physical store integration can have on multichannel retailing, influencing everything from consumer behavior to operational strategies and financial outcomes.

### 2.2.4 Consumer behavior and return management

The impact of store openings in omnichannel retail on return rates is influenced by various factors, including consumer behavior, operational strategies, and the interaction between physical and online retail channels. Studies have shown that the concept of "showrooming," where consumers browse in physical stores but make purchases online, significantly affects the profits of brick-and-mortar retailers. This highlights the interconnected nature of retail channels and the importance of understanding the effects of store openings on return rates in an omnichannel environment (Balakrishnan et al., 2014).

Furthermore, the presence of physical stores significantly affects consumer behavior at the point of purchase. Anderson et al. (2009) discuss the "option value of returns," suggesting that consumers perceive physical stores as a safer option due to easier return processes. This perception can lead to more thoughtful purchasing decisions and lower return rates. Li and Pan (2023) also highlight that return shipping fees are significant considerations in consumer purchase decisions, particularly in sectors like fashion where return rates are notably high. The clear visibility of return policies and costs at physical sales points can dramatically influence the likelihood of post-purchase returns.

Moreover, high return rates in online shopping impact retailers financially and shape consumer perceptions and behaviors. Li and Choudhury (2020) emphasize the need for proactive return management to address these challenges effectively. Gao and Su (2016) observe that physical showrooms not only decrease online returns but stimulate overall demand, indicating that the experiential aspects of physical shopping can positively affect online consumer behaviors.

### 2.2.5 Strategies for reducing return rates

Integrating physical stores with online platforms involves navigating a numerous logistical and operational challenges, especially in managing returns. The implementation of BORIS (buy online, return in-store) strategies significantly aids in this integration by allowing customers to return online purchases in physical stores. This approach not only enhances customer interactions but also streamlines return processes, contributing to improved operational efficiency and cost savings.

### 2.2.6 Conclusion

In summary, the integration of physical and online retail channels presents significant opportunities for addressing the challenges associated with high return rates, particularly in the context of omnichannel retailing. The literature emphasizes the critical role of physical stores in mitigating return rates by providing tactile product experiences, personalized customer service, and immediate resolution of issues, which collectively enhance consumer confidence and satisfaction. Moreover, strategic approaches such as enhanced product descriptions, sizing information, and the implementation of BORIS strategies are essential for reducing returns and improving operational efficiency. This comprehensive understanding of return rate behavior and effective management strategies not only fosters greater customer loyalty but also drives overall profitability. The evidence highlights that a well-coordinated omnichannel strategy, leveraging the strengths of both physical and online retail environments, is crucial for optimizing returns management and achieving sustainable retail success.

## 2.3 Conceptual framework

This section develops a conceptual framework to understand how changes in the return fraction (i.e., the fraction of revenues that are returned) impact the net revenue uplift and how this, in turn, is influenced by changes in the return rate, average price per unit sold, and average price per item returned. This framework aims to provide a structured approach to analyze the relationships between components, and their effects on a retailer's performance.

We will first explain the discrepancy between net revenue uplift and gross revenue uplift as a function of return fraction uplift, followed by the decomposition of return fraction uplift and its relationship with return rate uplift. Next, we will discuss how return rate uplift influences the divergence between net revenue and gross revenue uplifts. Finally, we will present a conceptual model summarizing these relationships. Table 2.1 introduces the notation used in the equations used in this section. For the derivation of the expressions, we refer the reader to Appendix B.

**Table 2.1** Definitions of variables

| **Variable** | **Definition** |
|---|---|
| $GR_{before}$ | the gross revenues before store opening |
| $GR_{after}$ | the gross revenues after store opening |
| $RF_{before}$ | fraction of gross revenues returned (return fraction) before store |
| $Uplift_{GR}$ | the uplift of gross revenues due to store opening |
| $Uplift_{NR}$ | the uplift of net revenues due to store opening |
| $Uplift_{RR}$ | the uplift of return revenues due to store opening |
| $Uplift_{AS}$ | the uplift of the average price per item sold due to store opening |
| $Uplift_{AR}$ | the uplift of the average price per returned item due to store opening |

### 2.3.1 Discrepancy between net revenue and gross revenue uplift

The change in return behavior after a store opening determines how the net revenue uplift deviates from the gross revenue uplift. This can be expressed as:

$$Uplift_{NR} = Uplift_{GR} - \underbrace{\left(\frac{RF_{before}}{1 - RF_{before}}\right)}_{\substack{\text{(a) The odds ratio of} \\ \text{the return fraction} \\ \text{before store opening}}} . \underbrace{(Uplift_{RR} - Uplift_{GR})}_{\substack{\text{(b) The excess uplift of} \\ \text{return revenues} \\ \text{over gross revenues}}} \tag{2.1}$$

This expression indicates that when return revenues increase more than gross revenues, the net revenue uplift is reduced by an additive factor. This factor is the product of (a) the odds ratio of the return fraction before store opening and (b) the excess uplift of return revenues over gross revenues. In other words, the degree to which the net revenue uplift is lower than the gross revenue uplift depends on the degree to which the increase in return revenues outpaces the increase in gross revenues, and this effect is more pronounced when the fraction of revenues that are returned was larger prior to the store opening. Conversely, the net revenue uplift could exceed the gross revenue uplift if the store opening results in a higher growth of gross revenues compared to return revenues.

When return revenues grow more (or less) than gross revenues, this implies a positive (or negative) uplift in the return fraction. This relationship can be reformulated as:

$$Uplift_{RR} - Uplift_{GR} = Uplift_{RF} \cdot (Uplift_{GR} + 1) \tag{2.2}$$

Combining the above equations result in:

$$Uplift_{NR} = (1 - C_{GR\to NR}) \cdot Uplift_{GR} - C_{GR\to NR} \tag{2.3}$$

where

$$C_{GR\to NR} = \left(\frac{RF_{before}}{1 - RF_{before}}\right) \cdot Uplift_{RF} \tag{2.4}$$

In other words, a higher return fraction uplift decreases the net revenue uplift both additively and multiplicatively, while a lower return fraction uplift increases it using the same mechanism.

$$\begin{array}{lll} Uplift_{NR} < Uplift_{GR} & \text{when } C_{GR\to NR} > 0 & \Rightarrow Uplift_{RF} > 0 \\ Uplift_{NR} = Uplift_{GR} & \text{when } C_{GR\to NR} = 0 & \Rightarrow Uplift_{RF} = 0 \\ Uplift_{NR} > Uplift_{GR} & \text{when } C_{GR\to NR} < 0 & \Rightarrow Uplift_{RF} < 0 \end{array} \tag{2.5}$$

### 2.3.2 Decomposition of the return fraction uplift

Next, we decompose the drivers of the return fraction uplift. The derivation of the expressions can be found in the Section B3 of Appendix B.

$$Uplift_{RF} = C_{RRate\to RF} \,.\, (1 + Uplift_{RRate}) - 1 \tag{2.6}$$

$$\text{,where } C_{RRate\to RF} = \frac{Uplift_{AR}+1}{Uplift_{AS}+1}$$

In this equation, $Uplift_{AR}$ is the uplift in the average price per returned item, and $Uplift_{AS}$ is the uplift in the average price per item sold. Therefore,

$$Uplift_{RF} < Uplift_{RRate} \quad \text{when } C_{RRate\to RF} < 1 \Longrightarrow Uplift_{AR} < Uplift_{AS} \tag{2.7}$$
$$Uplift_{RF} = Uplift_{RRate} \quad \text{when } C_{RRate\to RF} = 1 \Longrightarrow Uplift_{AR} = Uplift_{AS}$$
$$Uplift_{RF} > Uplift_{RRate} \quad \text{when } C_{RRate\to RF} > 1 \Longrightarrow Uplift_{AR} > Uplift_{AS}$$

The return fraction uplift equals the return rate uplift when the average prices of sold and returned items change similarly. Otherwise, the return fraction uplift is larger (smaller) than the return rate uplift when the uplift in the average price per returned item exceeds (or is less than) that of the average price of a sold item.

### 2.3.3 Explaining discrepancy between net revenue and gross revenue uplift as a function of return rate uplift

The formulation in Equation (2.8) explains how the net revenue uplift deviates from the gross revenue uplift as a function of return rate uplift. Here, $C_{GR\to NR}$ is an adjustment factor that encapsulates the influence of the odd ratio for return fraction before the store opening ($RF_{before}$), the uplift in the average price per returned item ($Uplift_{AR}$), the uplift in the average price per item sold ($Uplift_{AS}$), and the return rate uplift ($Uplift_{RRate}$). This factor indicates that the net revenue uplift is reduced by an amount dependent on the extent to which the return rates increase relative to the gross revenues.

$$Uplift_{NR} = [1 - C_{GR\to NR}]\, Uplift_{GR} - C_{GR\to NR} \tag{2.8}$$

where

$$C_{GR\to NR} = \left(\frac{RF_{before}}{1 - RF_{before}}\right) . Uplift_{RF}$$

$$Uplift_{RF} = (C_{RRate\to RF} \,.\, (1 + Uplift_{RRate}) - 1)$$

$$C_{RRate\to RF} = \frac{Uplift_{AR} + 1}{Uplift_{AS} + 1}$$

Specifically, when the return rate uplift and the changes in the average prices of sold and returned items are substantial, they exert a significant downward pressure on the net revenue uplift. Conversely, if the return rate uplift is controlled and the prices of sold and returned items are managed effectively, the net revenue uplift can approach or even exceed the gross revenue uplift.

### 2.3.4 Explaining divergence in uplift of return fraction and return rate

The difference between net revenue uplift ($Uplift_{NR}$) and gross revenue uplift ($Uplift_{GR}$) can be understood by examining how the return fraction ($Uplift_{RF}$) and return rate ($Uplift_{RRate}$) interact with the changes in the average price per returned item ($Uplift_{AR}$) and the average price per sold item ($Uplift_{AS}$).

The equations (2.9) below summarize the conditions under which $Uplift_{NR}$ is less than $Uplift_{GR}$.

(I) $Uplift_{NR} < Uplift_{GR}$ and $Uplift_{AR} < Uplift_{AS}$ when $0 < Uplift_{RF} < Uplift_{RRate}$ $\Rightarrow 0 < \frac{Uplift_{RF}}{Uplift_{RRate}} < 1$ (2.9)

(II) $Uplift_{NR} < Uplift_{GR}$ and $Uplift_{AR} = Uplift_{AS}$ when $0 < Uplift_{RF} = Uplift_{RRate}$ $\Rightarrow \frac{Uplift_{RF}}{Uplift_{RRate}} = 1$

(III) $Uplift_{NR} < Uplift_{GR}$ and $Uplift_{AR} > Uplift_{AS}$ when $0 < Uplift_{RRate} < Uplift_{RF}$ $\Rightarrow 1 < \frac{Uplift_{RF}}{Uplift_{RRate}}$

The equations (2.10) below summarize the conditions under which $Uplift_{NR}$ is greater than $Uplift_{GR}$.

(IV) $Uplift_{NR} > Uplift_{GR}$ and $Uplift_{AR} < Uplift_{AS}$ when $Uplift_{RF} < Uplift_{RRate} < 0$ $\Rightarrow \frac{Uplift_{RF}}{Uplift_{RRate}} > 1$ (2.10)

(V) $Uplift_{NR} > Uplift_{GR}$ and $Uplift_{AR} = Uplift_{AS}$ when $Uplift_{RF} = Uplift_{RRate} < 0$ $\Rightarrow \frac{Uplift_{RF}}{Uplift_{RRate}} = 1$

(VI) $Uplift_{NR} > Uplift_{GR}$ and $Uplift_{AR} > Uplift_{AS}$ when $Uplift_{RRate} < Uplift_{RF} < 0$ $\Rightarrow 0 < \frac{Uplift_{RF}}{Uplift_{RRate}} < 1$

***Dampening Effect* (I and VI).** When $C_{RRate \to RF} < 1$, the return fraction uplift is lower than the return rate uplift. This indicates a muted impact, where the return rate uplift has a reduced effect. This condition suggests that the average price per sold item increases more than the average price per returned item, which is favorable for maintaining a higher net revenue uplift. Note that this also includes cases where $C_{RRate \to RF} < 0$, indicating a further reduction in impact.

The primary focus of the *Dampening Effect* is on increasing the average selling price and maintaining low return rates to enhance net revenue uplift. Retailers should implement pricing strategies that increase the average selling price of items, such as offering value-added services, introducing premium product lines, or running effective marketing campaigns. Additionally, they can improve product quality to justify higher prices and reduce the likelihood of returns. Enhancing customer satisfaction and providing accurate product descriptions can ensure that customers are satisfied with their purchases, thereby reducing return rates. Targeted promotions can also be used to sell higher-margin items, contributing to a higher average price.

***Amplifying Effect*** **(III and VI).** When $C_{RRate \to RF} > 1$, the return rate uplift is amplified such that the return fraction uplift is even higher. This suggests that efforts to manage return rates are critical. The higher return rates, particularly when the average price per returned item increases more than the price per sold item, indicate that higher-value items are being returned more frequently, necessitating targeted strategies to address this issue.

The primary focus of the *Amplifying Effect* is on actively managing and reducing return rates, particularly for high-value items, to avoid financial losses. Retailers should implement strategies to manage and reduce return rates, such as stricter return policies or incentives for exchanges rather than returns. Educating customers through detailed product information and virtual try-on tools can help them make informed purchasing decisions, decreasing the likelihood of returns. Providing excellent post-purchase support and customer service can address issues that might otherwise lead to returns. Furthermore, using data analytics to identify patterns in returns can enable businesses to develop targeted interventions to address the root causes.

***Balanced Impact*** **(II and V).** When $C_{RRate \to RF} = 1$, the return fraction uplift and the return rate uplift are equal, highlighting a neutral scenario where the changes in average prices do not disproportionately affect the return fraction. This balance ensures that the net revenue uplift closely follows the gross revenue uplift, assuming other factors remain constant. This applies to both positive (condition II) and negative (condition V) uplifts.

For retailers, this means maintaining stable pricing strategies to ensure that neither returns nor sales disproportionately impact overall revenue. Retailers can focus on consistent pricing and quality strategies that maintain equilibrium between sales and returns, ensuring steady revenue streams without significant fluctuations due to returns.

Overall, these conditions—*Dampening Effect*, *Amplifying Effect*, and *Balanced Impact*—provide a comprehensive framework for retailers to strategically address both revenue enhancement and cost reduction to achieve optimal financial performance. While the *Dampening Effect* emphasizes boosting the revenue side through higher average prices and fewer returns, the *Amplifying Effect* focuses on minimizing the costs associated with returns, especially of high-value items. By understanding and managing the return rates and return fractions, retailers can tailor their strategies to maximize net revenue uplift.

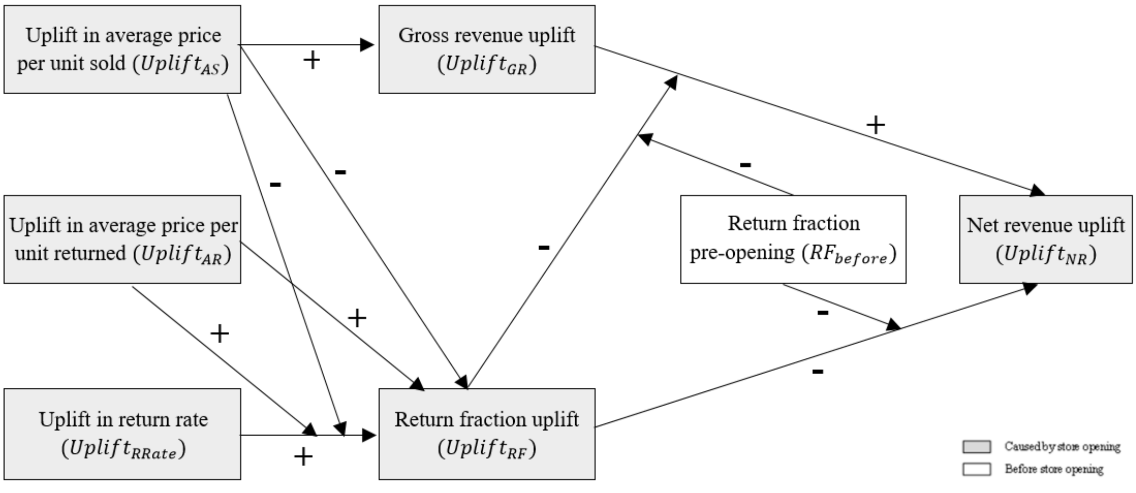
### 2.3.5 Conceptual model

The conceptual framework in Figure 2.1 summarizes (a) how the return fraction uplift causes the net revenue uplift to deviate from the gross revenue uplift, and (b) how the return rate uplift, average price per unit sold, and average price per item returned jointly affect the return fraction uplift. We refer the reader to the Appendix B for a full list of variables and their definitions used in this framework.

The net revenue uplift ($Uplift_{NR}$) is influenced by the gross revenue uplift ($Uplift_{GR}$) and adjusted by the return fraction uplift ($Uplift_{RF}$). Gross revenue uplift directly contributes to the net revenue uplift but can be diminished by a high return fraction uplift. A positive return fraction uplift indicates that more revenues are being returned, which reduces the net revenue uplift.

The uplift in the return rate ($Uplift_{RRate}$) directly impacts the return fraction uplift. An increase in return rate typically leads to a higher return fraction uplift unless offset by changes in average prices. Higher prices per unit sold ($Uplift_{AS}$) can lead to increased gross revenues, which may influence the return fraction if the return rate does not change proportionally. Similarly, an increase in the average price per unit returned ($Uplift_{AR}$) may lead to a higher return fraction uplift, particularly if the return rate remains constant or increases.

**Figure 2.1** Conceptual framework on how return uplift affects net revenue uplift




The return fraction uplift is influenced by the uplift in return rate and the changes in average prices per unit sold and returned. Positive uplift in the average price per unit sold leads to a positive gross revenue uplift. Higher return rates and higher prices of returned items generally increase the return fraction uplift. A positive return fraction uplift reduces the net revenue uplift, highlighting the importance of managing return behaviors post store opening.

The framework encapsulates how the net revenue uplift ($Uplift_{NR}$) is derived from the gross revenue uplift ($Uplift_{GR}$) and how it is influenced by the return fraction uplift ($Uplift_{RF}$). It emphasizes the following:

- A positive gross revenue uplift generally leads to a positive net revenue uplift unless significantly offset by the return fraction uplift.
- The return fraction uplift moderates the impact of gross revenue uplift on net revenue uplift. Managing return rates, especially for higher priced items, and the prices of sold and returned items is crucial for optimizing net revenue uplift.
- Changes in return rate and the average prices of sold and returned items collectively determine the return fraction uplift, which in turn affects the net revenue uplift.

This comprehensive framework provides a structured approach to analyze and understand the relationships between sales, returns, and revenues, offering a basis for empirical validation and strategic decision-making.

## 2.4 Research setting and data

This study leverages the same research setting and data from the online-first consumer electronics retailer in Western Europe, as detailed in Chapter 1. The retailer, known for its advanced omnichannel strategies, expanded into offline retail, providing an ideal environment to examine the impact of physical stores on return behaviors. At the beginning of the study period, the retailer operated a website, mobile app, and eleven brick-and-mortar stores in two neighboring countries. This setting allows us to explore how different store formats and omnichannel services influence return patterns. The collaboration benefits from the retailer's mature omnichannel integration, offering services such as reserve-online, pickup-in-store, ship-from-store, ship-to-store, and buy-online, return-in-store. These services enable a seamless customer experience across channels, which is crucial for understanding return behaviors in an omnichannel context.

Our data set includes detailed transaction records linked to specific customers, including information on stock-keeping units (SKUs), number of units sold, prices, and return instances. Customer identification relies on various identifiers (e.g., phone number, email, customer account), ensuring accurate tracking of purchase and return activities. To analyze the data, we aggregated it at the month-zipcode level, using zipcode characteristics obtained from the National Statistics Bureau, which include socio-demographics, urbanization levels, and access to transportation.

As shown in Table 2.2, the mean and standard deviation of each variable before store opening are compared for all stores, control zip codes, and treated zip codes. The table also includes the difference between treated and control zip codes, along with the p-value to indicate statistical significance. The results show no statistically significant difference between the treatment and control groups.

**Table 2.2** Pre-opening zip code characteristics

| Variables | All stores | Control zip codes | Treatment zip codes | Treatment vs. Control | p value t test |
|---|---|---|---|---|---|
| Return Revenues (EUR) | 15.65<br>(23.58) | 16.50<br>(24.26) | 14.50<br>(22.56) | -2.00<br>(1.33) | 0.133 |
| Return Fraction (%) | 2.98<br>(4.32) | 3.01<br>(4.22) | 2.94<br>(4.46) | -0.07<br>(0.11) | 0.538 |
| Return Rate (%) | 3.31<br>(3.53) | 3.27<br>(3.36) | 3.36<br>(3.76) | 0.09<br>(0.09) | 0.338 |

Notes.
Standard deviations are reported in parentheses.
*p<0.05, **p<0.01, ***p<0.001

## 2.5 Research design and model specification

### 2.5.1 Research design

Our research design and model specification closely follow the methods described in Chapter 1, with necessary adaptations to focus on return behaviors. We use a difference-in-differences (DiD) approach combined with propensity score matching to estimate the causal impact of physical store presence on return rates and related metrics. The identification strategy capitalizes on the staggered opening of three new stores, comparing affected zip codes (within a 20-minute travel radius of the new stores) with unaffected ones. This approach allows us to isolate the effect of store openings on return behaviors while controlling for other factors.

Using the same identification strategy as in Chapter 1, we identified 192 out of 2,388 zip codes as affected by the store openings. Promotion, shipping, and return policies were consistent across all zip codes, verified with our retail partner. We employed a DiD specification to compare changes in return rates and related variables between affected and unaffected zip codes. To address potential confounding factors, we applied propensity score matching to construct control groups from unaffected zip codes. This method, detailed in Chapter 1, involved logistic regression to predict the probability of a zip code becoming part of a store region, based on observed characteristics. Matching was performed using nearest-neighbor algorithms within a specified caliper, ensuring similarity between treated and control groups.

Given the staggered introduction of stores, we used an event-style DiD approach to account for the timing differences. This method, as described by Sun and Abraham (2021), addresses biases in traditional TWFE DiD models by providing unbiased estimates in cases of staggered adoption. The relative time variable was binned into three-month intervals to ease interpretation and to account for the staggered nature of store openings. This approach allows us to distinguish between short- and long-term effects of store openings on return behaviors.

### 2.5.2 Model specifications

Next, we describe the analyses required to estimate the return fraction uplift and the return rate uplift at the store level. We also consider the uplifts in average price per unit sold and average price per unit returned.

*Modeling the return fraction and return rate uplifts.* The return fraction is the proportion of gross revenues that are returned. It is calculated as $RF_{it} = \frac{RR_{it}}{GR_{it}}$ where $RR_{it}$

is the return revenues for store $i$ time $t$ and $GR_{it}$ is the gross revenues for store $i$ at time $t$. The return rate ($RRate_{it}$) is the proportion of units sold that are returned, calculated as $RRate_{it} = \frac{UR_{it}}{US_{it}}$ where $UR_{it}$ is the number of units returned for store $i$ at time $t$ and $US_{it}$ is the number of units sold for store $i$ at time $t$. Modeling fractions is challenging as there are many observations that take a value of 0 when there are no returns or a value of 1 when all items are returned.

$$\log (RF_{it}) = \log \frac{RR_{it}}{GR_{it}} = \log\ (RR_{it}) - \log(GR_{it}) \qquad (2.11)$$

$$\log (RR_{it}) = \log \frac{UR_{it}}{US_{it}} = \log\ (UR_{it}) - \log(US_{it})$$

According to Villadsen and Wulff (2021), one approach is to use the log odds ratio and fractional regression with a logit link function. The $logit$ function transforms a proportion (e.g. return fraction) into an unbound continuous variable, defined as $logit(p) = log\ \left(\frac{p}{1-p}\right)$. Here $p$ represents either the return fraction $RF_{it}$ or the return rate $RRate_{it}$. The model is specified as follows:

$$\text{logit}\ (DV_{it}) = \log\left(\frac{DV_{it}}{1 - DV_{it}}\right) = \lambda_i + \mu_t + \sum_{s=1}^{3}\sum_{\substack{k=-8 \\ k\neq -1}}^{1} \alpha_{s(i),k} \cdot OPEN_{s,k} + \varepsilon_{it} \qquad (2.12)$$

, where

$$OPEN_{s,k} \ = \ I[i\ \in \text{store region of } s] \cdot I[relative\ time\ bin_{ts} = k]$$

The dependent variable $DV_{it}$ represents either the return fraction $RF_{it}$ or the return rate $RRate_{it}$. The model incorporates fixed effects for each store $\lambda_i$ and each time period $\mu_t$ to control for unobserved heterogeneity. The indicator function $OPEN_{s,k}$ equals 1 if zip code iii is within the store region $s$ and relative time bin $k$, and 0 otherwise, ensuring that the impact of store openings is accurately captured. Finally, the error term $\epsilon_{it}$ accounts for random variations in the model. The uplift effect $\alpha_{s(i),k}$ quantifies the effect of the store opening on the dependent variable within the specified time bin.

## 2.6 Results

This section presents the short-term impacts of the openings of large stores on returns. Specifically, the analysis focuses on the net revenue uplift ($Uplift_{NR}$), gross revenue uplift ($Uplift_{GR}$), and return revenue uplift ($Uplift_{RR}$), as well as the return fraction uplift ($Uplift_{RF}$) and return rate uplift ($Uplift_{RRate}$).

***Impact on net and gross revenues.*** In Chapter 1, we showed that $Uplift_{NR}$ was consistently less than $Uplift_{GR}$ for both Large Store 1 and Large Store 2. Specifically, Large Store 1 experienced a net revenue uplift of 22.78% compared to a gross revenue uplift of 23.54%, while Large Store 2 saw a net revenue uplift of 20.62% versus a gross revenue uplift of 21.18%. Units sold increased significantly for Large Store 1 (24.35%, 95% CI [15.96, 32.74]) and Large Store 2 (20.03%, 95% CI [13.64, 26.41]). These findings indicate that while there were significant increases in gross revenues and units sold post store opening, the net revenue gains were lower than gross revenues.

***Impact on return revenues.*** The return revenue uplift ($Uplift_{RR}$) was significantly higher for both stores. Large Store 1 exhibited an uplift of 141.75% (95% CI [30.04, 253.45]), and Large Store 2 demonstrated a 96.52% uplift (95% CI [11.81, 181.22]). Similarly, the number of total units returned also registered significant increases: Large Store 1 saw a 36.18% increase (95% CI [19.88, 52.49]), and Large Store 2 experienced a 27.60% increase (95% CI [13.68, 41.52]).

***Discrepancy between net revenue and gross revenue uplift.*** According to our conceptual framework, $Uplift_{NR}$ is less than $Uplift_{GR}$ when the return fraction uplift ($Uplift_{RF}$) is positive. This situation arises when return revenues increase at a faster rate than gross revenues. The observed increase in return revenues which outpaced gross revenues, suggests a positive shift in the return fraction, consistent with Equation (2.2). This shift indicates a reduction in net revenue uplift, moderated by an additive factor derived from (a) the pre-opening odds ratio of the return fraction and (b) the differential uplift between return and gross revenues.

Based on our empirical data, the return fraction is estimated to be 2.76% and 2.86% for Large Store 1 and 2, respectively. After removing outliers, these rates slightly decreased.[1] The differential uplift between return and gross revenues can be estimated by subtracting the gross revenue uplift from the return revenue uplift. For Large Store 1, this

---

[1] Specifically, we calculated the lower and upper thresholds for outliers based on 1.5 times the interquartile range for the return fraction before opening. Outliers were then identified and removed from the dataset. Following this data cleaning process, we recalculated the return fraction before store openings for each of the large stores. The return fraction is estimate to be 2.07% and 2.16% for Large Store 1 and 2, respectively. Please refer to Appendix B for evidence showing no statistically significant difference between the treatment and control groups for each large store. Thus, the pre-opening odds ratio of the return fraction is approximately 2% to 3% for both stores.

yields a differential uplift of 118.21% (141.75% - 23.54%). Similarly, for Large Store 2, the differential uplift is 75.34% (96.52% - 21.18%).

For Large Store 1, this additive factor is -2.5% (based on an odds ratio of 2.11%).[2] For Large Store 2, it is approximately -1.66% (based on an odds ratio of 2.21%). The estimated discrepancy between net revenue uplift and gross revenue uplift ($Uplift_{NR} - Uplift_{GR}$) from our empirical analysis was -0.76% for Large Store 1 and -0.56% for Large Store 2. While the additive factor and the estimated discrepancy share the same direction, their magnitudes differ suggesting the presence of estimation bias. However, the framework provides valuable insights into the direction and magnitude of these relationships. Table 2.3 below summarizes the comparison of predictions for Equation (2.1) from the conceptual framework and estimated values from our empirical analysis for Large Store 1 and 2.

**Table 2.3** Comparison of predictions from Equation (2.1) and empirical estimations

| | **Large store 1** | **Large store 2** |
|---|---|---|
| (a) odds ratio $\left(\frac{RF_{before}}{1-RF_{before}}\right)$ | 2.11 | 2.21 |
| (b) $Uplift_{RR} - Uplift_{GR}$ | 118.21 | 75.34 |
| additive factor - (a) . (b) | (2.50) | (1.66) |
| $Uplift_{NR} - Uplift_{GR}$ | (0.76) | (0.56) |

***Impact on return fraction and return rate.*** The effects of store openings on return behaviors, measured through return fraction and return rate, are detailed in Table 2.4. The logit-transformed coefficients indicate the impacts at Large Store 1 and Large Store 2. For Large Store 1, the coefficient for the return fraction is 0.158, indicating no significant change post-opening, while for Large Store 2, the coefficient is 0.142, also indicating no significant change.

Conversely, the return rate presents a different scenario. For Large Store 1, the return rate coefficient is 0.249, signifying a notable increase post-opening, while for Large Store 2, the coefficient is 0.196, reflecting a similar trend. These statistically significant positive coefficients for the return rate in both stores indicate an increase in the odds of items being returned following the store openings. Specifically, the estimated effect sizes for the return rate uplifts, $Uplift_{RRate}$, are 27.11% for Large Store 1 and 20.81% for Large

[2] The additive factor estimates of -3.4% for Large Store 1 and -2.22% for Large Store 2 are based on the odds ratios without removing outliers.

Store 2. Please see Appendix B for the calculations on return rate uplift. These findings illustrate the importance of effective return rate management to mitigate its impact on net revenue uplift, as outlined in the conceptual framework.

**Table 2.4** Results for logit (Return Fraction) and logit (Return Return)

| | Large store 1 | | Large store 2 | |
|---|---|---|---|---|
| Dependent Variable | Coef. | [95% CI] | Coef. | [95% CI] |
| $logit\ (RF_{it})$ | 0.158<br>(0.16) | [-0.16, 0.48] | 0.142<br>(0.13) | [-0.11, 0.40] |
| $logit\ (RRate_{it})$ | 0.249*<br>(0.11) | [0.03, 0.47] | 0.196*<br>(0.10) | [0.00, 0.40] |

Notes.
Standard deviations are reported in parentheses. *p<0.05, **p<0.01, ***p<0.001

***Explaining divergence in uplift of return fraction and return rate.*** Our empirical analysis revealed that $Uplift_{NR}$ was less than $Uplift_{GR}$, consistent with our conceptual framework which anticipated a positive $Uplift_{RF}$. We can derive the predicted $Uplift_{RF}$ from Equation (2.2) using the ratio of the differential uplift between return and gross revenues ($Uplift_{RR} - Uplift_{GR}$) to ($Uplift_{GR}$) + 1. The return fraction is predicted to increase by 95.7% for Large Store 1 and 62.2% for Large Store 2 following the store openings. This indicates that the predicted $Uplift_{RF}$ is higher than the estimated $Uplift_{RRate}$.

Further, utilizing Equation (2.6), and considering the predicted $Uplift_{RF}$ alongside the estimated $Uplift_{RRate}$, we computed the value of $C_{RRate \to RF}$ as approximately 1.54 for Large Store 1 and 1.34 for Large Store 2. Our conceptual framework also posited that $C_{RRate \to RF}$ would be greater than 1 if the $Uplift_{RF}$ was higher than the $Uplift_{RRate}$, a condition that was met according to our findings. Consequently, this implies that the uplift in the average price per returned item ($Uplift_{AR}$) was higher than the uplift in the average price per item sold ($Uplift_{AS}$), aligning with our expectations and emphasizing the interplay between various revenue components in the context of retail store openings.

Our findings are consistent with the amplified effect, specifically condition III, as outlined in the conceptual framework. This condition suggests that when the uplift in the average price per returned item ($Uplift_{AR}$) is higher than the uplift in the average price per item sold ($Uplift_{AS}$), the impact of the return rate uplift ($Uplift_{RRate}$) on the return fraction uplift ($Uplift_{RF}$) is amplified. This indicates that the higher average price of returned items compared to sold items intensified the effect of increased return rates, resulting in a more significant uplift in the return fraction. These findings show the importance of managing high-value returns to mitigate their impact on net revenue uplift, as higher-value returns can disproportionately amplify the return fraction, thus reducing net revenue gains despite increases in gross revenue.

Next, we estimate from our data the uplift for average price per item sold and average price per returned item.

***Impact on average price per item sold and returned item.*** The analysis of uplifts in average prices per sold and returned items reveals distinct patterns for Large Store 1 and Large Store 2 following the store openings, though the findings are not statistically significant. For Large Store 1, the average price per sold item experienced a minor decline of 1.4%, while the average price per returned item saw a more significant reduction of 8.2%. This indicates that, in Large Store 1, the items returned post-opening tended to be of lower value compared to those sold. Conversely, Large Store 2 exhibited a slight increase of 0.6% in the average price per sold item and a more substantial increase of 3.7% in the average price per returned item. This suggests that in Large Store 2, higher-value items were returned more frequently post-opening. Overall, these findings highlight that while Large Store 2 saw a higher uplift in the average price per returned item compared to the average price per sold item, Large Store 1 experienced a greater decline in the average price per returned item relative to the average price per sold item. However, it is important to note that these results are not statistically significant, indicating that these patterns could be due to random variation rather than systematic effects of the store openings.

## 2.7 Discussion

This study provides insights into the impact of store openings on return rates within an omnichannel retail environment. Key findings highlight a significant increase in return rates following store openings, suggesting that operational strategies implemented at the time of store launches substantially influence return behaviors.

The return fraction uplift was higher than the uplift of item return rates, indicating that the uplift in the average price per returned item was higher than the uplift in the average price per item sold. This amplification effect suggests that the higher average price of returned items compared to sold items intensified the effect of increased return rates, resulting in a more significant uplift in the return fraction.

Our findings have important implications for retailers, particularly in managing return rates, and interventions targeted at high-value items. Effective returns management is critical; retailers need to diligently ensure that high-ticket items have a lower return rate by taking proactive measures. This includes providing detailed product explanations, offering robust post-purchase support, and enhancing customer education to minimize returns. By focusing on these strategies, retailers can mitigate the negative impacts on net revenue uplift. Implementing measures that curb returns for expensive items can stabilize revenue streams and maintain consistent profit margins despite fluctuations in return rates.

Moreover, controlling the return rate is crucial. Enhancing product descriptions, boosting customer satisfaction, and providing robust post-purchase support can directly influence the return fraction uplift and, consequently, the net revenue uplift. Detailed and accurate product descriptions help set correct customer expectations, reducing return likelihood. Excellent post-purchase support, including straightforward return processes and responsive customer service, is vital in lowering return rates. Additionally, offering incentives for retaining purchases or alternatives such as exchanges can further manage the return rate, leading to enhanced customer satisfaction and improved financial performance.

The study also highlights the importance of designing targeted interventions for high-value items, which are more likely to be returned and have a significant impact on net revenue due to their higher price points. Strategies such as personalized customer support and detailed product information can be highly effective. Implementing innovative solutions like virtual try-ons and predictive models that proactively address potential return issues can lead to successful interventions. These targeted strategies ensure that the negative impacts on net revenues are minimized, and high-value items positively contribute to the retailer's overall financial health.

However, it is crucial to acknowledge the limitations of this study concerning the verification of the conceptual framework. The absence of significant empirical findings for the return fraction, average price per sold item, and average price per returned item suggests that the study does not fully validate the underlying conceptual framework. Future research should revisit this framework in a different setting to gain more definitive insights into these relationships. Additionally, it would be beneficial to investigate the source of the significant increase in return rate. Is it due to products purchased online being returned in-store, or are store purchases more likely to be returned than online purchases? This counterintuitive aspect could yield interesting insights and warrants further investigation. Further studies could also benefit from conducting category-level analyses to explore how different product types are impacted by the identified factors, allowing for a more customized approach to managing returns across various retail categories.

Chapter 3

# RIGHTSIZING STORE LABOR: A FIELD EXPERIMENT

## 3.1 Introduction

Retailers prioritize delivering high-quality customer service by maintaining sufficient staffing levels and employing skilled, knowledgeable staff. Adequate staffing and expertise are essential for converting customer traffic into sales and fostering repeat business, ultimately enhancing customer satisfaction and loyalty. Effective labor management is thus crucial for optimizing sales and maintaining a competitive edge. However, payroll costs, typically around 10% of sales, represent a significant variable expense for stores, second only to the cost of goods sold (Kesavan & Mani, 2015). Prof. Marshall Fisher at The Wharton School of the University of Pennsylvania notes that while reducing headcount saves payroll, it can harm the top line (Lee, 2023). This underscores the importance of balancing staffing costs with service levels. Overstaffing leads to wasted resources, while understaffing compromises service quality and customer satisfaction.

Retailers face the challenge of striking the right balance between customer-centric measures that enhance service effectiveness and operational efficiency-driven measures that control costs. The key lies in matching store staffing levels with fluctuating customer traffic to maintain service standards, reduce wait times, and manage payroll expenses efficiently. However, this is complicated by the imperfect nature of the data retailers have to use at the time of scheduling store labor.

Traditionally, employee schedules are made in advance, based on estimates of expected customer traffic or sales forecasts. This approach aims to align labor supply with anticipated demand but often falls short due to the inherent variability and unpredictability of customer traffic. Despite efforts to develop accurate demand forecasts and robust scheduling systems, imbalances between planned and required labor capacity persist. Factors such as forecast inaccuracies and employee absenteeism contribute to these discrepancies, resulting in overstaffing or labor shortages on the day of service (Mani et al., 2015).

In the Netherlands, regulations require schedules to be known weeks in advance, allowing employees to plan their lives around them. This is also important in a tight labor market where retaining talent is crucial. Maintaining advanced schedules not only complies with legal requirements but also enhances employee satisfaction and retention by providing predictability and stability in their work-life balance.

Responsible scheduling practices, such as publishing schedules well in advance and maintaining predictability, are essential for both employees and companies. These practices lead to increased employee well-being, reduced stress, and improved job satisfaction, which in turn enhance productivity and overall company performance. However, real-time adjustments can also be managed responsibly by ensuring they are voluntary. Just-in-time scheduling, as discussed by Kamalahmadi et al. (2021), involves adjusting staffing levels in real-time to meet fluctuating customer demand. This strategy enhances operational flexibility and aligns labor supply with actual customer traffic, potentially reducing labor costs and improving service levels. However, it introduces unpredictability in work schedules, affecting employee well-being and performance.

To balance flexibility with stability, it is crucial to consider legal and logistical constraints, such as labor laws and employee contracts. Fair workweek laws mandate predictable and stable scheduling practices to protect employee rights and well-being (Lambert et al., 2014; Schneider & Harknett, 2019). These laws require advance notice of schedules, employee consent for changes, and predictability pay for last-minute adjustments, ensuring control over schedules and adequate rest between shifts. Responsible scheduling practices that provide consistent, predictable, and controllable work schedules are crucial for enhancing employee well-being. Kesavan et al. (2022) demonstrate that such practices can lead to increased employee effort, reduced absenteeism, and improved store performance. Emphasizing predictable and responsible scheduling practices is essential for maintaining employee morale and job satisfaction. The management of our retail partner commented that this is especially important in the tight labor market that retailers face nowadays. Improving scheduling predictability and stability can enhance store productivity and sales without increasing labor hours.

Unpredictable and unstable work schedules negatively impact worker well-being, leading to reduced job satisfaction and increased stress. Addressing these issues through responsible scheduling practices improves employee outcomes and store performance

(Lambert et al., 2014; Schneider & Harknett, 2019). Implementing scheduling practices that balance business needs with employee well-being can enhance store execution and productivity. Even with advanced scheduling systems and responsible practices, the inherent unpredictability of customer demand poses challenges. Real-time staffing adjustments offer a promising solution by providing on-the-fly adaptability to labor management, thus correcting forecast inaccuracies. While these adjustments increase operational flexibility, they must be voluntary to ensure employee well-being and productivity are maintained.

This study introduces the concept of real-time staffing adjustments in retail operations, implemented on a voluntary basis per the retailer's request. This approach provides significant flexibility to meet fluctuating customer demand while allowing employees to adjust their hours during their shifts according to their preferences and availability. Such flexibility not only aligns labor supply with actual demand but also contributes to employees' work-life balance, reducing stress and improving job satisfaction. For instance, employees can make extra income by opting for additional hours during peak times or leave early when they have personal commitments like a birthday party or good weather to enjoy. This ability to tailor work hours to personal needs enhances overall well-being.

By benefiting employees, this approach enhances overall productivity, as satisfied employees are typically more engaged and efficient. Real-time adjustments enable retailers to take corrective measures in labor planning with the consent of employees. By allowing stores to dynamically adjust staffing levels based on current store conditions, retailers can better match labor supply to actual customer traffic. This process involves monitoring real-time data from queue management systems and point-of-sale systems to make informed decisions about staffing needs. Addressing discrepancies between planned and actual demand helps maintain optimal staffing levels, thus enhancing store performance.

In this chapter, we examine the impact of granting stores the autonomy to adjust real-time staffing levels on store labor productivity. Specifically, we investigate how the direction of real-time staffing adjustments (increasing or decreasing staffing levels) affects store productivity. We report results from a field experiment at a consumer electronics retailer that rolled out the flexibility to adjust staffing levels in real-time based on current store conditions. We persuaded the retailer to intermittently block the ability to adjust hours on and off for a few stores during our experiment, enabling us to assess the effects on performance. We monitored and captured the adjustments and the reasons behind them.

Our study is the first to examine real-time staffing adjustments in the consumer electronics retail sector, which has unique staffing needs compared to other retail sectors like apparel. The integration of insights from data sheets filled out by shift leads highlights the need for real-time corrections, providing a detailed account of the reasons for scaling labor hours. There is a significant gap in the existing literature on dynamic staffing strategies specifically within the retail sector. While much research has focused on static scheduling and just-in-time practices in other service industries, there is limited empirical evidence on the effectiveness of voluntary and real-time staffing adjustments in retail. Our study addresses this gap by providing unique insights into the implementation and impact of real-time staffing adjustments in consumer electronics retail. The rigorous empirical methodologies we use, such as the field experiment design, provide robust causal evidence on the effectiveness of real-time staffing adjustments. Our comprehensive data integration from multiple sources, including point-of-sale data, queue management systems, and labor data, allows for a detailed analysis of staffing adjustments and their impact on store performance.

The introduction of real-time staffing adjustments led to a statistically significant 6.24% increase in store productivity for stores granted the autonomy to adjust staffing levels in real-time. This intent-to-treat (ITT) estimate is conservative, as it includes all stores assigned to the intervention, regardless of whether they utilized the staffing adjustments. Further analysis, measured as Average Treatment Effect on the Treated (ATT) also revealed a positive causal impact on store productivity. Specifically, when stores opted to upscale labor hours, productivity improved significantly by 17.45%, as increased sales during high visitor traffic outweighed the added labor hours, demonstrating the effectiveness route.

While the overall labor hours did not show a significant change, there was a noticeable trend towards more efficient use of monthly planned hours. The intervention led to scaling activity half of the time, with downscaling occurring two-thirds of the time and upscaling one-third of the time. This balance indicates that stores were able to dynamically adjust their workforce to meet real-time demand fluctuations, optimizing their monthly planned labor hours and enhancing labor efficiency.

The intervention did not significantly affect sales, suggesting that while operational efficiency improved, the direct impact on sales may require a longer observation period or additional supportive measures. The study explored two routes to productivity improvements: upscaling during high traffic for effectiveness and downscaling during low traffic for efficiency. However, the effectiveness route, with significant productivity boosts through upscaling, proved to be more impactful.

Overall, the study demonstrates that real-time staffing adjustments are an effective tool for improving store productivity. Granting stores the flexibility to make these adjustments enhances operational efficiency and leads to better overall performance. These insights can inform retailers' labor management strategies and contribute to developing more adaptive and employee-friendly scheduling practices.

The rest of this chapter is organized as follows. Section 3.2 reviews the literature on staffing practices and identifies research gaps. Section 3.3 presents our conceptual framework on real-time staffing adjustments. Section 3.4 describes our data sources and variables. Section 3.5 outlines our field experiment design, intervention strategy, and data collection methods. Section 3.6 discusses our findings, including intent-to-treat and average treatment effect results. Section 3.7 concludes with managerial implications, limitations, and future research opportunities. Additional data and results are in Appendix C.

## 3.2 Literature review

Our study builds upon a comprehensive body of research examining staffing practices within service industries. Empirical investigations in diverse sectors—including healthcare, as highlighted by Peng et al. (2023) and Chan et al. (2017), alongside the restaurant industry, as studied by Kamalahmadi et al. (2021) and Tan and Netessine (2014)—emphasize the importance of optimal staffing. These studies demonstrate that precisely calibrated staffing levels are crucial for enhancing service quality and operational efficiency. These implications are directly transferable to the retail sector, particularly for high-service retailers where maintaining exceptional service quality is a core proposition.

### 3.2.1 Impact of staffing levels on retail performance

Empirical evidence robustly confirms the significant impact of staffing levels on key retail performance metrics, such as sales volume, conversion rates, basket sizes, and overall productivity. A clear correlation indicates that staffing aligned with customer demand substantially improves these outcomes, thereby emphasizing the critical role of dynamic staffing strategies in retail performance enhancement.

Recent scholarly contributions, particularly those of Lee et al. (2021), Musalem et al. (2021), Chuang et al. (2016), and Kesavan et al. (2014), explore the relationship between staffing decisions and retail performance. These studies explain how changes in staffing levels affect important metrics such as sales volume and customer satisfaction. While Lee et al. (2021) and Musalem et al. (2021) explore strategies around customer assistance and efficient backend operations, Chuang et al. (2016) propose labor-planning frameworks designed to maximize profitability amid variable customer traffic. Kesavan et al. (2014) advocate for a balanced labor management strategy, cautioning against excessive reliance on flexible labor resources.

Further investigations by Mani et al. (2015), Perdikaki et al. (2012), and Netessine et al. (2010) provide a detailed examination of how labor alignment with traffic patterns affects store productivity and sales, offering a comprehensive understanding of operational management. Foundational research by Ton (2009) and Ton and Huckman (2008) discusses how labor significantly influences profitability through enhanced service quality. Similarly, Fisher et al. (2006) identify critical drivers of retail store performance, particularly the roles of inventory and staffing levels.

Our study aims to extend these discussions by examining the impact of real-time schedule adjustments on retail store performance, with a particular focus on productivity and sales. By employing rigorous empirical methodologies, we seek to offer new insights into the effective implementation of dynamic staffing strategies, thereby enhancing operational flexibility and efficiency. This exploration is expected to contribute substantially to the existing literature, providing practical strategies for retail managers to optimize workforce deployment in response to evolving market conditions.

### 3.2.2 Aligning staffing with customer demand

Operational flexibility, a cornerstone of service operations, revolves around staffing strategies that adapt to fluctuating customer demands. This flexibility is achieved through various means, including cross-training employees to enhance throughput (Gans et al., 2003; Iravani et al., 2005), employing part-time and temporary workers (Kesavan et al., 2014), and implementing just-in-time work schedules that allow for capacity shifting across different time periods (Kamalahmadi et al., 2021).

However, the implementation of real-time schedule adjustments poses unique challenges and opportunities. Just-in-time (JIT) scheduling, including short-notice and real-time schedules, has been extensively studied in restaurant settings (Kamalahmadi et al., 2021). Short-notice scheduling asks employees a couple of days in advance to work extra shifts, while real-time scheduling asks them to extend shifts on the day of service. While short-notice scheduling has minimal productivity impact, real-time scheduling can reduce productivity by up to 4.4% due to fewer opportunities for up-selling and cross-selling. This reduction occurs because employees may work more slowly or take longer breaks, avoiding customer interactions as a stress-coping mechanism (Krischer et al., 2010). Such behaviors can lead to perceived poor service quality and decreased sales without significantly affecting meal duration (Kamalahmadi et al., 2021). Additionally, research suggests that shifting towards more predictable scheduling patterns could significantly increase profitability.

Unpredictable scheduling practices, characterized by erratic work schedules, can have profound implications for worker well-being and productivity. Studies have highlighted how such schedules adversely affect employees, leading to reduced job satisfaction and overall well-being (Henly & Lambert, 2014; Kesavan et al., 2022; Lu et al., 2022). Thus, there is a growing emphasis on the need for predictable and flexible scheduling practices to enhance both employee satisfaction and operational efficiency.

### 3.2.3 Voluntary real-time adjustment

Our study focuses on the impact of real-time voluntary staffing adjustments in a retail setting, addressing the need for operational flexibility amidst high demand variability while overcoming the unpredictability of just-in-time scheduling practices. Unlike mandatory adjustments, these voluntary adjustments allow employees to opt-in based on their availability and preferences. Thompson (1999) was a pioneer in formally discussing real-time schedule adjustments involving last-minute changes to workforce schedules to align with actual customer demand and operational conditions, particularly relevant in high-volume operations like the retail sector where high customer variability enhances the ability to predict business volume early in the day.

Real-time schedule adjustments have been extensively studied across various sectors, demonstrating broad applicability and a significant impact on operational performance. Studies have proposed mathematical formulations for real-time staffing adjustments based on updated forecasts. Hur et al. (2004) investigated adjustment decisions in quick service restaurants, showing that heuristic-based real-time adjustments made by managers frequently surpass static scheduling methods in terms of profitability and customer service levels. Easton and Goodale (2005) introduced a model that addresses the effects of unplanned absenteeism through real-time schedule adjustments. Their research links staffing policies with operational flexibility, illustrating that effective real-time adjustments can significantly mitigate the negative impacts of absenteeism on productivity. Mehrotra et al. (2010) developed a dynamic model for call centers that adjusts staffing levels based on real-time demand forecasts. Their findings indicate that incorporating updated demand predictions can reduce staffing costs and improve service levels by better aligning resources with actual demand. Mac-Vicar et al. (2017) presented empirical results from a retail context, where their algorithmic approach to real-time adjustments led to an 18% reduction in lost profits due to staffing inefficiencies. Their model illustrates the potential for real-time adjustments to enhance both profitability and employee satisfaction by minimizing unnecessary schedule disruptions.

Our work contributes to this body of research, presenting the first field experiment, to our knowledge, to study the impact of real-time schedule adjustments in a retail setting. Through detailed empirical analysis, we aim to provide valuable insights into the effectiveness of real-time voluntary staffing strategies, ultimately guiding organizations towards more efficient and sustainable workforce management practices.

## 3.3 Conceptual framework

Real-time adjustments in labor hours can enhance store productivity through two primary mechanisms, as depicted in the conceptual model (Figure 3.1). There are two distinct routes: upscaling for effectiveness when traffic is much higher than anticipated, and downscaling for efficiency when traffic is much lower than anticipated. Each route offers a different pathway to productivity improvements.

In our study, we define store productivity as store-level labor productivity, measured by sales ($) per labor hour, a common operationalization in retail research (Kesavan et al., 2022). This metric allows us to assess how well stores convert labor hours into sales revenue, providing a clear indicator of efficiency and effectiveness in workforce management.

***Effectiveness (Marketing) Route.*** Increasing labor hours during periods of high customer traffic can lead to sales that are elastic to labor, meaning the increase in sales is proportionally larger than the increase in labor hours (elasticity > 1). Perdikaki et al.

(2012) found that an increase in store labor is associated with better sales performance. Mani et al. (2015) demonstrated that eliminating understaffing can lead to substantial increases in sales and profitability. Similarly, Ton (2009) noted that higher labor allocation improves profitability by enhancing conformance quality, which is defined as adherence to company processes and standards. Fisher et al. (2006) emphasized that customer satisfaction positively impacts sales per labor hour in retail stores.

During high customer traffic, well-staffed stores can provide faster service, reduce customer wait times, and enhance customer experience, ultimately driving higher sales (Parasuraman et al., 1988). This aligns with Ton (2009), who found that increasing labor and reducing the workload per employee can decrease errors or corner-cutting, thus maintaining quality and potentially enhancing productivity. Thus, upscaling labor hours during peak times can directly translate to increased sales and productivity.

***Efficiency (Operations) Route.*** Conversely, reducing labor hours during periods of lower than expected customer traffic can improve productivity by cutting costs without significantly impacting sales. This happens when the decrease in sales is proportionally smaller than the reduction in labor hours. This concept aligns with lean management and operational efficiency literature, which suggests that optimizing labor during off-peak times maintains service levels while reducing operational costs (Womack & Jones, 1996). Chuang et al. (2016) emphasized the importance of effective management of store labor in successful retail operations, as labor performs crucial service-related and production-like tasks.

However, decreasing labor hours in retail stores may not necessarily lead to improved store productivity and operational efficiency. Ton (2009) found that increasing labor and reducing employee workload can decrease errors and maintain quality, which can enhance productivity. Additionally, Chapados et al. (2014) highlighted the critical importance of effective sales staff scheduling for profitable retail operations, indicating that workforce deployment efficiencies translate into margin expansion and improved profitability. Basker (2005) mentioned that increasing average efficiency in the retail sector, potentially by reducing the number of workers needed per sale, could be a consequence of certain strategies like Wal-Mart's entry into the market. However, this does not directly address the impact of decreasing labor hours on store productivity and operational efficiency.

Therefore, while some efficiency gains may be achieved through optimizing labor allocation, reducing labor hours without careful consideration of tasks and workload may not necessarily lead to improved store productivity and operational efficiency.

The conceptual model (Figure 3.1) illustrates both routes, defining productivity as the ratio of sales to labor hours. The arrows represent the expected directions, with "+" indicating a positive relationship and "−" indicating a negative relationship. The effectiveness route is expected to have a more positive impact on sales than hours, whereas the efficiency route focuses on optimizing labor without significant sales loss.

The primary challenge in real-time scheduling lies in choosing between upscaling for immediate sales gains and downscaling for cost efficiency. Upscaling involves allocating more labor hours during peak traffic periods to enhance customer service and maximize sales. This strategy requires higher labor costs, with the critical task of ensuring that the additional sales revenue justifies these costs. Retailers must balance the potential benefits of increased sales against the risk of overstaffing and reduced profitability if customer traffic does not meet expectations.

On the other hand, downscaling focuses on reducing labor hours during low traffic periods to cut costs while maintaining sales levels. The risk is that excessive staff reduction might negatively impact customer service and satisfaction, potentially leading to long-term losses in customer loyalty and sales. Retailers must balance the immediate cost savings with the need to maintain a satisfactory customer experience.

The inherent tension arises from the risks associated with making real-time adjustments. It can be tempting for retailers to maintain the status quo rather than risk making changes. Downscaling often appears less risky and easier to justify to headquarters, as it directly cuts costs. However, the real challenge is in making informed decisions that strike a balance between the immediate financial benefits of cost-cutting and the long-term advantages of investing in customer service and satisfaction through upscaling. Based on these mechanisms, we propose the following hypotheses:

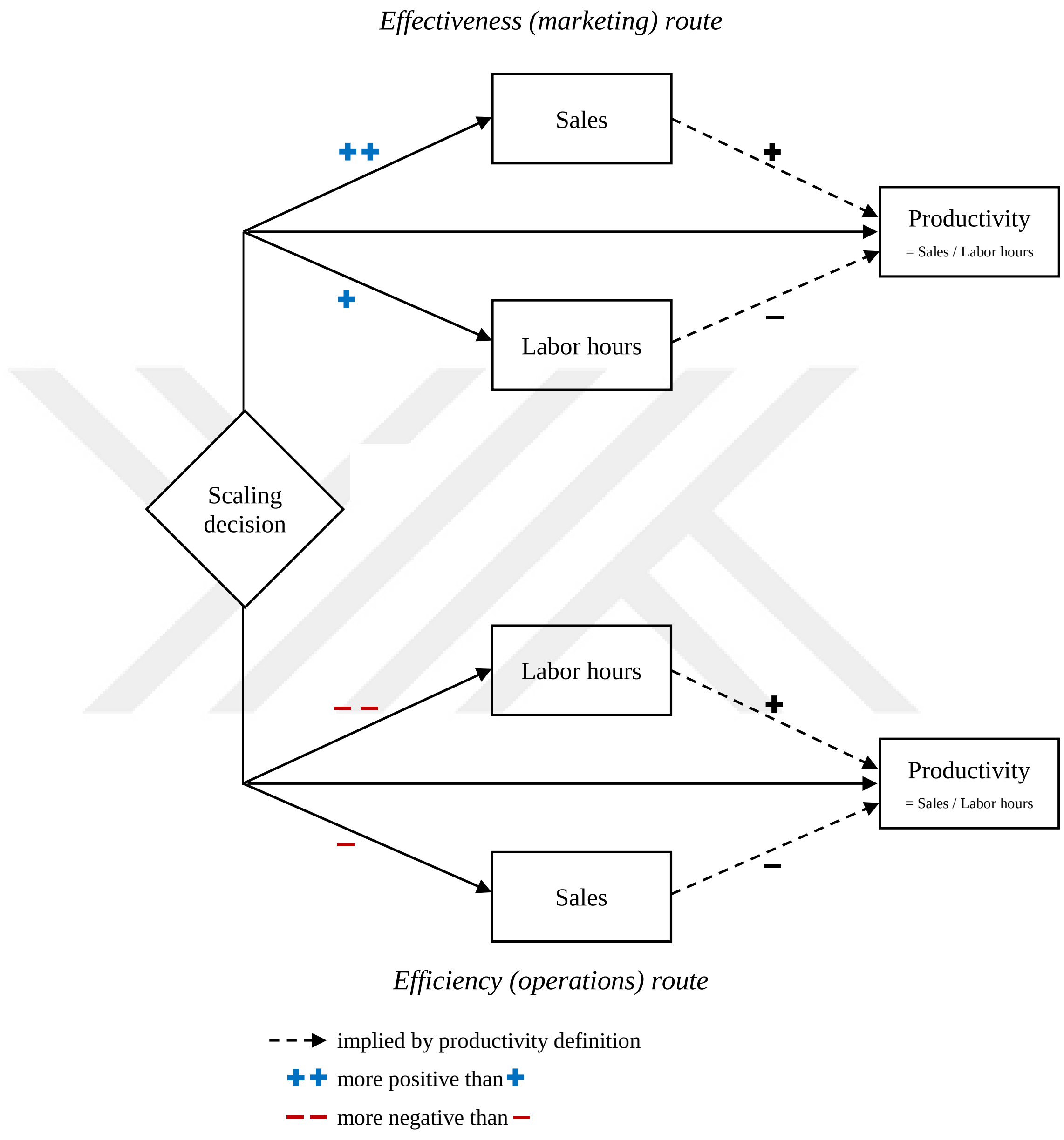
**Hypothesis 1.** Upscaling labor hours during high visitor traffic increases productivity by boosting sales more than the increase in labor hours.

This hypothesis is based on the premise that additional labor during peak times directly translates to higher sales.

**Hypothesis 2.** Downscaling labor hours during low visitor traffic increases productivity by maintaining sales while reducing labor hours.

This hypothesis posits that labor reductions during off-peak times can lower costs without significantly affecting sales.

**Figure 3.1** Conceptual framework




## 3.4 Data and variable definitions

We next present our data sources, followed by a description of the variables (dependent, independent, and controls) we use in our study.

### 3.4.1 Data sources

We acquire the following data for all 12 stores during the study period from November 2022 to July 2023 from the retailer:

*Point-of-Sale Data.* This includes order-level information on all store transactions, whether originated at the store or picked up at the store. It provides insights into retail performance and consumer behavior, such as the number of transactions, store sales volume and attach rate.

*Queueing Management System Data.* The retailer utilized queueing management systems with counters installed at the entrances of the stores to facilitate customer flow and service efficiency. Upon entering the store, customers are greeted by the counter, which prompts them to take a number and select the purpose of their visit from options such as advice, service, pick-up, or repair. Additional instructions are provided based on the selected option; for example, customers selecting "pick-up" are directed to the designated counter at the back of the store, with a prompt indicating they will be served when it's their turn. This dataset captures valuable information on visitor arrival time, purpose of visit, service start and completion times, and the employee who assisted the customer. It allows for a comprehensive analysis of customer behavior and flow within the store, enabling us to derive metrics such as the number of visitors, served versus no-show customers, time spent serving customers, average service time, and average wait time. While visitor forecast data is available at the store-hour level for labor planning, we complement this information with actual hourly visitor counts derived from the Queueing Management System. This dual approach provides detailed insights into both expected and observed customer traffic patterns, facilitating our analysis of store productivity.

*Labor Data.* This dataset includes planned and worked employee hours, providing insights into staffing levels and resource allocation within the store. Planned hours are provided at the hourly level with associated roster activity planned, while worked hours are aggregated at the employee-date level.

*Employment Data.* This dataset contains information on employees' job titles and dates they were hired and promoted to other positions. It helps us compute tenure at the current position and the year of employment with the company.

*Real-time Staffing Adjustment Data.* This dataset provides detailed information on real-time staffing adjustments made by shift leads, enabling us to evaluate their impact on store performance. It captures whether staffing levels were scaled up or down, the number of labor hours adjusted, and the identity of the shift lead. Additionally, it documents the motivation behind each decision, the factors considered, and the real-time data metrics used. The dataset also includes feedback from shift leads for future adjustments and evaluations of the impact of previous adjustments on store performance.

### 3.4.2 Variables

First, we define the dependent variables, explanatory variables, and control variables in our main analysis and then provide summary statistics for all these variables.

*Dependent variables.* Our primary objective is to examine the effect of real-time scheduling on store productivity. Store productivity ($PROD_{it}$) is measured as the dollar sales generated per labor hour ($SALES_{it}$ / $LBR_{it}$) in store $i$ on day $t$. To understand this impact further, we analyze its components, including sales ($SALES_{it}$) and labor hours $LBR_{it}$ in store $i$ on day $t$. Sales may be influenced by factors such as basket value ($BV_{it}$), which represents the ratio of sales volume to the number of transactions in store $i$ on day $t$ ($SALES_{it}$ / $TRANS_{it}$). Additionally, we consider the conversion rate ($CR_{it}$), calculated as the ratio of transactions to tickets issued in store $i$ on day $t$ ($TRANS_{it}$ / $TICK_{it}$).

*Key variables of interest.* Our primary focus is on the net real-time schedule adjustments made to labor hours aimed at optimizing store productivity. We define the key variable of interest as Scaling ($SCAL_{it}$), representing the net adjustment in labor hours on day $t$ in store $i$. This variable is computed as the sum of upscaling and downscaling. Upscaling ($UPSCAL_{it}$) denotes the increase in labor hours to meet higher demand, while downscaling ($DOWNSCAL_{it}$) involves reducing labor hours during periods of lower demand. These variables offer valuable insights into the adaptability of stores in adjusting their staffing levels to match fluctuating customer traffic. The equation $SCAL_{it} = UPSCAL_{it} - DOWNSCAL_{it}$ is the algebraic sum of $UPSCAL_{it}$ and $DOWNSCAL_{it}$, representing the net adjustment.

*Control variable.* In our study, we utilize $TICK_{it}$ as a control variable, representing the number of customer tickets obtained from queueing counters installed at the store entrances. This metric serves as a proxy for store visitors, as customers are required to obtain a ticket upon entering the store to interact with a customer service employee.

The choice of $TICK_{it}$ as a control variable is crucial for several reasons. Firstly, it accurately measures customer engagement, distinguishing active customers from mere browsers, and controls for daily fluctuations in store traffic. This allows us to isolate the effects of staffing adjustments on productivity, sales, and labor hours. Moreover, $TICK_{it}$

reduces bias from non-engaged visitors and enhances the robustness of our econometric model. By excluding individuals who do not impact sales or require staff assistance, we obtain more precise estimates of the intervention's impact on store productivity and operational efficiency. By using $TICK_{it}$ as a control variable, we aim to control for variations in visitor footfall and focus our analysis on active customer interactions. This approach enhances the internal validity of our findings by ensuring that observed changes in performance metrics are not confounded by variations in general store traffic.

Table 3.1 describes the variables that we use, and Table 3.2 pre-experiment summary provides the descriptive statistics for all variables and present the Pearson correlation coefficients among all variables, respectively.

**Table 3.1** Variable definition

| Variable | Description |
|---|---|
| Productivity ($PROD_{it}$) | The ratio of $SALES_{it}$ to $LBR_{it}$ in store $i$ on day $t$. |
| Labor hours ($LBR_{it}$) | The total number of labor hours worked in store $i$ on day $t$. |
| Sales ($SALES_{it}$) | The total dollar value of sales transactions that occur in store $i$ on day $t$. |
| Conversion rate ($CR_{it}$) | The ratio of $TRANS_{it}$ to $TICK_{it}$ in store $i$ on date $t$. |
| Number of transactions ($TRANS_{it}$) | The total number of transactions that occur in store $i$ on day $t$. |
| Tickets ($TICK_{it}$) | The number of customer tickets issued throughout the day from queueing counters installed in front of the stores, including all types, in store $i$ on day $t$. |
| Basket value ($BV_{it}$) | The ratio of $SALES_{it}$ to $TRANS_{it}$ in store $i$ on date $t$. |
| Scaling ($SCAL_{it}$) | The net adjustment in labor hours in store $i$ on day $t$, calculated as the sum of upscaling and downscaling. |
| Upscaling ($UPSCAL_{it}$) | The increase in labor hours in store $i$ on day $t$, reflecting additional staffing needs. |
| Downscaling ($DOWNSCAL_{it}$) | The reduction in labor hours in store $i$ on day $t$, indicating a decrease in staffing levels. |

The pre-experiment summary statistics provide valuable insights into retail performance metrics. Productivity ($PROD_{it}$) is strongly correlated with sales ($SALES_{it}$), number of transactions ($TRANS_{it}$), and tickets ($TICK_{it}$), indicating that higher sales and transaction counts boost productivity. Labor hours ($LBR_{it}$) show a weak positive correlation with productivity, suggesting that merely increasing labor hours does not significantly enhance productivity. Additionally, labor hours have a weak positive correlation with the conversion rate ($CR_{it}$) but a negative correlation with basket value ($BV_{it}$), implying that more labor hours do not necessarily translate to higher basket value. This highlights the need for strategic labor deployment rather than just increasing staffing. Furthermore, the conversion rate's moderate correlation with basket value and productivity shows the importance of effective customer conversion strategies to drive sales and productivity.

**Table 3.2** Pre-experiment summary statistics of the variables and Pearson Correlation Coefficients

| Variable name | Mean | SD | Median | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|---|---|---|
| (1) Productivity ($PROD_{it}$) | 396.95 | 148.32 | 376.61 | 1.00 | | | | | | |
| (2) Labor hours ($LBR_{it}$) | 111.92 | 40.28 | 104.50 | **0.15** | 1.00 | | | | | |
| (3) Sales volume ($SALES_{it}$) | 45,308.42 | 25,949.60 | 38,304.00 | **0.74** | **0.74** | 1.00 | | | | |
| (4) Number of transactions ($TRANS_{it}$) | 151.53 | 73.24 | 132.00 | **0.60** | **0.78** | **0.93** | 1.00 | | | |
| (5) Conversion rate ($CR_{it}$) | 0.70 | 0.09 | 0.70 | 0.04 | **-0.06** | 0.03 | **0.15** | 1.00 | | |
| (6) Basket value ($BV_{it}$) | 292.23 | 65.09 | 289.32 | **0.63** | **0.22** | **0.52** | **0.22** | **-0.30** | 1.00 | |
| (7) Tickets ($TICK_{it}$) | 217.01 | 100.96 | 193.00 | **0.59** | **0.81** | **0.93** | **0.96** | **-0.09** | **0.30** | 1.00 |

*Notes.* Number of observations is 2,241. Bold denotes statistical significance at the 1% level. SD, standard deviation.

## 3.5 Experiment design and implementation

We partnered with the same consumer electronics retailer described in Chapters 1 and 2, which initially focused on online sales but has since expanded to include offline retail. During our experiment, the retailer operated through its website, mobile app, and a growing network of 22 brick-and-mortar stores across three neighboring countries, with plans to expand to 50 stores by the end of next year. These stores offer customers hands-on product experiences, immediate access to online orders, and personalized advice from expert staff. Store personnel assist with product usage, repairs, and returns, while the retailer integrates digital features into its offline stores. For instance, the mobile app includes an in-store mode that adds customers to the queue upon arrival and provides access to online product information. Additionally, self-service queueing kiosks allow customers to register and receive a ticket for their place in line, ensuring minimal waiting times and keeping customers informed.

The store staff comprises various roles, including managers, shift leads, service associates, warehouse, and repair employees. Store managers work full-time, and while the majority of shift leads also have full-time roles, scheduling may vary for other staff members. Store managers play a pivotal role in overseeing operations and are expected to actively engage on the sales floor. Similarly, shift leads and service associates primarily operate on the shop floor, directly interacting with customers to ensure exceptional service. However, all staff members are flexibility and may assist in customer service duties regardless of their designated roles.

The retailer follows a structured scheduling process facilitated by the workforce planning team from the the corporate office. Based on the traffic forecasts for each store location across specific dates, the workforce planning team allocates shifts to employees based on their basic schedules, which encompass the minimum contracted hours. Employee preferences, such as specific availability on certain days, are taken into account during this process.

Employees do not dedicate all their working hours to assisting customers. They must also allocate time for various other tasks, including cleaning, breaks, and unspecified duties. To accommodate this, an occupancy rate is established, indicating the proportion of scheduling hours allocated for customer assistance versus other store activities. The capacity planning team highlights the inherent limitations associated with a fixed occupancy rate, as such rates may vary significantly across different days and store locations. Additionally, employees are entitled to breaks, the scheduling of which is managed directly by the shift lead on the day itself.

Preliminary staffing schedules, comprising both assigned and available shifts for future selection, undergo a collaborative review with the store manager to ensure alignment with store needs. Within a four-week window preceding the schedule's commencement, employees have the opportunity to select shifts corresponding to their remaining contractual hours, with the option to secure additional hours if desired. As the schedule's execution date approaches, typically with two weeks remaining, the staff schedule is finalized and published.

*Shift Lead position.* In November 2022, our retail partner introduced the role of Shift Lead for the Day (SL), designating them as a "primus inter pares" employee responsible for positioning employees on the floor, monitoring performance, adjusting roles (e.g., assigning someone to the front desk when not actively selling), deciding on breaks, and stimulating non-selling tasks during quieter periods. Please find a detailed job description for SLs in the Appendix C1.

To prepare SLs for their responsibilities, a comprehensive training program was held from October 31 to November 18, 2022. The training aimed to equip SLs with the skills to make informed day plans, present impactful day starts, and manage daily operations effectively. Preparation steps included reviewing the previous day's handover, checking emails, and creating the day plan by assigning employees and scheduling breaks. Shift leads were taught to ensure each team member knows their role and responsibilities. Training also covered managing daily operations by monitoring waiting and transaction times, positioning the team effectively, and managing breaks. Scenarios included managing the pick-up area, addressing employee health issues, handling busy periods, and managing personal breaks. Key learnings included being prepared for unexpected situations, recognizing the positive impact of effective presentation, understanding the central role of the shift lead, and the importance of effective communication. The training provided guidance on managing daily store operations, focusing on planning, communication, and adaptability, ensuring shift leads were equipped to handle various challenges efficiently. The same instructor delivered this training across all stores, ensuring consistent content and no differences in training across different stores and regions.

*Shift Lead dashboard.* The introduction of the Shift Lead Dashboard in March 2023 significantly bolstered operational insights by providing near real-time metrics encompassing store traffic, service levels, sales, and productivity[3]. A comprehensive list of these metrics is available in the Appendix Section C1. Shift Leads (SLs) are subsequently entrusted with the responsibility of monitoring (near) real-time sales per hour worked and customer waiting time to effectively guide employees on the sales floor. This dashboard acts as a centralized hub, aggregating data from customer tickets and transactions, thereby equipping Shift Leads with the necessary information to make well-informed staffing decisions. It's important to note that SLs may exhibit varying levels of prior experience within the retailer, having served as store employees before their promotion to the SL position, which can result in differing contextual understandings and decision-making approaches among them.

*Real-time staffing adjustments.* In May 2023, our retailer partner recognized the challenge of aligning staffing levels with fluctuating customer demands and granted Shift Leads the autonomy to adjust staffing throughout the day to optimize sales per hour worked. This autonomy, coupled with ongoing dashboard monitoring and business evaluation, allows Shift Leads to approach store associates and offer opportunities to adjust their scheduled hours voluntarily based on store needs. This means that Shift Leads can ask employees if they would like to leave early or stay late depending on the business requirements for that day.

Shift Leads use the criterion of "maximizing the sales per hour worked by adjusting intraday staffing up or down" to decide on staffing adjustments. Although the main

---

[3] Please refer to Appendix C1 for the figure illustrating the timeline of events for this study.

outcome variable is well specified, which could induce extra effort from Shift Leads and employees (Locke & Latham, 2002, 2006), Shift Leads are not provided with the “optimal” level of sales per hour worked or precise instructions on how to achieve this. There are no specific guidelines given to Shift Leads on how to communicate the need for volunteers. Some Shift Leads post notices on the employee communication board, while others approach each employee individually to ask if they are willing to adjust their hours. Shift Leads are authorized to make real-time adjustments such as increasing labor hours by asking an already working employee to voluntarily stay late or decreasing labor hours by allowing an employee to voluntarily leave early. They can also reallocate tasks based on current store conditions and demands. Specific training for making real-time staffing adjustments is planned to include interpreting the SL dashboard and using it to manage employees on the floor effectively. However, by the start of our experiment, this training had not yet commenced.

### 3.5.1 Intervention design

Initially, the implementation of real-time staffing adjustments was planned to be rolled out simultaneously across all stores. However, to evaluate the impact of these adjustments on store performance, we convinced the retailer to devise a field experiment with minimal alterations to the full rollout.

Our experimental design employed a unique approach by alternating the ability to make staffing adjustments on and off in three designated stores (referred to as "test stores") over a six-week period. Specifically, our intervention involved periodically disallowing Shift Leads from adjusting real-time staffing in these test stores. This method created variations both between and within stores concerning labor hours, enabling us to draw causal inferences regarding the effects of these real-time schedule adjustments on store productivity.

*Selection of stores.* Prior to the launch of the experiment, the retailer operated 14 stores in its country of origin, divided across three regions: North, West, and South, along with a few additional stores abroad.[4] Two stores in the North region were excluded from the analysis: one due to consistent understaffing and the other due to relocation to a larger premise during the experiment period. These exclusions were necessary to mitigate potential biases. Stores experiencing significant operational disruptions, such as

[4] The stores abroad had not yet seen the introduction of the shift lead role.

relocation and high vacancy rates, cannot effectively act on the autonomy of adjusting hours up or down due to these exogenous shocks. Including these stores would introduce noise and obscure the true effects of the intervention.

After these exclusions, we were left with 12 stores. We strategically selected three stores, one in each of the regions (North, West, and South), detailed in Table C1 in Appendix C. This selection ensured a balanced representation of different regions, and reduced the likelihood of shift leads communicating with each other across test stores. The selected stores were operationally stable, not recently opened, and size-wise representative, providing a reliable basis for the study and enhancing the robustness of the evaluation of the intervention's impact on store performance.

*Intermittent blocking.* We categorized the timeline into three periods as depicted in Figure 3.2: the "pre-experiment period," starting from November 7, 2022, with the introduction of the shift lead position; the "experiment period," commencing on May 15$^{th}$ upon granting real-time shift adjustment autonomy to shift leads; and the "post-experiment period," spanning two weeks following the experiment's conclusion. The experiment ended on June 25$^{th}$, and our data collection concluded on July 7$^{th}$, 2023, marking the end of the post-experiment period.

In our experimental implementation, we intermittently blocked the ability to scale staffing. At least one store was blocked each week, while at most two stores were blocked (never none nor all). Each store was shut down for half of the experiment duration (three out of six weeks), either independently or in conjunction with another test store. This unique design aimed to introduce variability in the on-off sequence, thereby generating greater variability compared to a simultaneous on-off intervention sequence across all stores.

Randomization was crucial for ensuring unbiased results. The order of test stores being turned on and off was randomized, further ensuring the randomness of the experimental design and minimizing potential biases. The on-off scheduling of staffing adjustments was randomized to prevent stores from anticipating their treatment status. Stores were not informed in advance when they would be in the "on" or "off" condition, minimizing potential behavior changes that could skew the results. We randomly determined the order of on-off scheduling and communicated this to the retail partner, who then informed store managers, maintaining the randomization integrity. This randomization ensured that any observed differences in productivity could be attributed to the intervention rather than other factors, enhancing the study's internal validity.

**Figure 3.2** Intervention timeline

| | Pre-experiment Nov 14 - May 14 | | | Experiment period May 15 - June 25 | | | | | | Post-experiment June 26 - July 9 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Week 1 | Week 2 | Week 3 | Week 4 | Week 5 | Week 6 | Week 7 | Week 8 |
| Test Store 1 | x | … | x | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| Test Store 2 | x | … | x | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 |
| Test Store 3 | x | … | x | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 |
| Store 4 | x | … | x | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Store 5 | x | … | x | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Store 6 | x | … | x | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| …….. | x | … | x | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Store 12 | x | … | x | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

*Notes.*
x indicates that the shift leads' autonomy to make intraday staffing changes has not rolled out yet.
1 indicates that shift leads have the autonomy to make intraday staffing changes.
0 indicates (=intervention) shift leads' autonomy to change intraday staffing is blocked.

### 3.5.2 Roll-out

On Friday, May 12th 2023, an email communication was disseminated to all stores to announce the launch of the experiment. This email provided an overview of the expectations from each store. Please refer to the Appendix C1 for the English translation of this communication. Notably, Shift Leads (SLs) were granted autonomy in making decisions about staffing adjustments, without explicit instructions on how to execute these decisions. For instance, SLs had the flexibility to determine how to communicate their scaling decisions and identify employees willing to adjust their hours voluntarily. This could include posting a note on the employee messaging board to inquire about voluntary hour adjustments.

To address concerns about the validity of staffing adjustments, clear schedules outlining when SLs could make staffing decisions were distributed to all stores prior to the intervention on Friday, May 12th. These schedules were made accessible through separate folders for test stores and control stores, ensuring clarity and accessibility for SLs. Throughout the 6-week experiment period, regional managers conducted regular weekly calls with store managers every Monday morning to remind them whether staffing adjustments were permitted for that week. This reminder was issued to all stores.

### 3.5.3 Monitoring and tracking real-time staffing adjustments.

To ensure meticulous oversight of real-time staffing adjustments, we implemented a hands-on approach by developing Google Sheets for stores to report their staffing changes. This initiative generated comprehensive and insightful data, empowering us to effectively monitor and track alterations as they occurred. It was mandatory for all Shift Leads (SLs) in both test and control stores to log and document their daily staffing adjustments. The Appendix C1 includes a detailed Real-time Staffing Adjustments Tracking Sheet template, complete with example lines outlining the reporting format and desired level of detail. This template prompts SLs to input various essential details, such as their name, scaling decision, time of the decision, reason for the adjustment, considerations made, metrics utilized, feedback for improvement, evaluation of previous decisions, and the number of hours scaled up or down.

Each form is stored separately to maintain data integrity, ensuring that Shift Leads (SLs) cannot access or modify spreadsheets from other stores. This system allowed our research team to monitor the sheets in (near) real time. Furthermore, we implemented a mechanism to promptly notify the retailer if any SLs failed to complete the form, guaranteeing the completeness and accuracy of our data collection process.

As SLs only had access to their respective store's form and could not view forms from other stores, our experiment design mitigated potential violations of SUTVA (Stable Unit Treatment Value Assumption). However, there might have been some in-store learning, as SLs within the same store could observe each other's actions.

Our research team conducted regular checks on the forms and alerted the retailer to any incomplete or conflicting information. The retailer then communicated with the respective shift leads to ensure completion or correction of their entries. Specific instances of non-compliance were rare, but when they did occur, they were promptly addressed through direct communication and additional guidance to the SLs involved. Data security and privacy were prioritized, with measures ensuring that all collected data was handled responsibly and securely.

## 3.6 Results

In this section, we present our models to explore the impact of granting stores autonomy to adjust real-time staffing levels on performance, sales and labor hours using daily store-level data. Our econometric analysis is conducted at the store-date level. We first introduce our model to examine the effect of granting this autonomy (intent-to-treat effect) and then discuss the impact of actual staffing adjustments made by stores (average treatment effect on the treated). We address potential endogeneity issues and our approach to mitigating biases. Finally, we present our findings.

### 3.6.1 Intent-to-treat effect

We estimate the impact of granting stores autonomy to adjust real-time staffing levels on performance by analyzing the intent-to-treat (ITT) effects. In this approach, each store's treatment condition is determined by assignment to the treatment group, regardless of the extent to which schedules were adjusted. Consequently, the ITT effects provide conservative estimates of the impact of making real-time schedule adjustments. The ITT effect is estimated using following specification:

$$\log(Y_{it}) = \delta\, TREAT_{it} + \beta\, log(TICK)_{it} + \alpha_i + \lambda_t + \varepsilon_{it}, \tag{3.1}$$

where $Y$ represents the outcome variable measuring performance (Productivity, Labor hours, Sales, Conversion Rate, Basket Value), the subscript $i$ denotes store, and $t$ denotes the date. The outcome variables are transformed into their natural logs to increase the normality of the residuals. The binary variable, $TREAT_{it}$ takes a value of 1 when shift leads of store $i$ have the autonomy to adjust real-time staffing levels on date $t$, and 0 otherwise. The coefficient of $TREAT_{it}$, $\delta$, captures the intervention effect. The components $\alpha_i$ and $\lambda_t$ reflect the unobserved store fixed effects and date x day dummies, respectively. The component $\varepsilon_{it}$ captures the idiosyncratic error that varies by store $i$ and time $t$. Standard errors are clustered at the store × week level, aligning with the level at which treatment is independently assigned (Roth et al., 2023).

Traffic is a well-known driver of sales in physical retail settings (Perdikaki et al., 2012). However, in our context, customers must obtain a ticket to receive assistance from floor staff. Therefore, we utilize the natural logarithm of daily tickets issued in store $i$ on date $t$, including both served and abandoned tickets, $\log(TICK)_{it}$, as the control in our model specification.

Table 3.3 presents the ITT model results for all outcomes. In column (1), the intervention effect (δ) on productivity is positive and statistically significant (0.061, p= 0.045), indicating a 6.24% increase (95% CI [6.18, 6.30]). This suggests that empowering shift leads to adjust staffing in real-time positively influences store productivity. For labor hours in column (2), although the coefficient is negative (-0.050), it is statistically insignificant, indicating no significant change in labor hours. However, the negative sign suggests a slight reduction in labor hours, consistent with the scaling activities conducted. Overall, scaling occurs half of the time, and when it does, it involves upscaling one-third of the time and downscaling two-thirds of the time.

Column (3) shows the effect on sales, with a positive but statistically insignificant coefficient (0.011). This result suggests a minimal impact on sales due to the intervention. Columns (4) and (5) represent the effects on conversion and basket value, respectively. Both coefficients (0.009 and 0.002) are positive but statistically insignificant, indicating little change in these metrics. Overall, the intervention led to a significant improvement in store productivity but had minimal effects on other key metrics such as sales, conversion, and basket value. This suggests that while real-time schedule adjustments may enhance operational efficiency, they may not necessarily translate into significant improvements in customer-facing metrics.

To put the 6.24% increase in productivity into context, we can compare it to productivity changes in other retail trade industries. Recent data by Bureau of Labor Statistics shows non-store retailers (primarily online businesses) had the highest productivity change at 11.1% in 2022 compared to the prior year. Miscellaneous store retailers saw a 7.9% increase, general merchandise stores 5.3%, and clothing and clothing accessory stores 3.6%. In comparison, electronics and appliance stores only observed a 2.5% increase, and health and personal care stores (e.g., pharmacies like CVS) observed a mere 0.3% increase. This highlights the significant potential of real-time staffing adjustments in enhancing productivity, as our observed increase surpasses that of several other retail sectors.

**Table 3.3** Intent-to-treat results of the model in (1)

| Variable | log(Productivity)<br>(1) | log(Labor hours)<br>(2) | log(Sales)<br>(3) | log(Conversion)<br>(4) | log(Basket value)<br>(5) |
|---|---|---|---|---|---|
| $TREAT$ | 0.061** | -0.050 | 0.011 | 0.009 | 0.002 |
| | (0.030) | (0.034) | (0.017) | (0.012) | (0.016) |
| Store fixed effects | Yes | Yes | Yes | Yes | Yes |
| Date x day of the week fixed effects | Yes | Yes | Yes | Yes | Yes |
| $N$ | 2911 | 2911 | 2911 | 2911 | 2911 |
| $R^2$ | 0.731 | 0.882 | 0.933 | 0.559 | 0.590 |

Notes. Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

To ensure the robustness of our findings, we conducted several robustness checks. We incorporated week fixed effects and day-of-the-week dummies, refined the definition of labor hours to include only those worked by floor staff and management, analyzed data for both the experiment period alone and the entire duration, excluding the 2-week post-experiment period. We excluded data from days when a store was closed due to a local holiday.

Additionally, we investigated whether the intermittent blocking intervention sequence introduced bias or cumulative learning by examining treatment effects on stores treated once versus those treated multiple times (on-off-on). While the intervention's positive impact on productivity remained consistent for stores treated once, stores treated twice showed statistically insignificant coefficients, suggesting no discernible difference in the intervention's impact. These findings, detailed in the Appendix C2, reinforce the reliability of our main results presented in Table 3.3.

### 3.6.2 Average treatment effect on the treated

To study the impact of scaling real-time staffing levels up or down on performance for stores with the autonomy to make adjustments, we define the variables $UPDIR_{it}$ and $DOWNDIR_{it}$. The variable $SCAL_{it}$, recorded in shift lead's data sheets, captures the cumulative net adjustment in labor hours resulting from real-time schedule adjustments. It is calculated as the sum of both upscaling ($UPSCAL_{it}$) and downscaling ($DOWNSCAL_{it}$) adjustments.

We characterize scaling activities in terms of the direction of labor hour adjustments. $UPDIR_{it}$ is an indicator variable taking a value of 1 if the cumulative net adjustment in labor hours ($SCAL_{it}$) is positive indicating that $UPSCAL_{it}$ (increasing labor hours) has prevailed over $DOWNSCAL_{it}$ for store $i$ on day $t$, or otherwise 0. Similarly, $DOWNDIR_{it}$ is an indicator variable taking a value of 1 if $SCAL_{it}$ is negative, indicating that $DOWNSCAL_{it}$ (decrease in labor hours) has prevailed over $UPSCAL_{it}$ for store $i$ on day $t$, or otherwise 0.

To estimate the effect of scaling up or down, conditional on being granted the autonomy to scale, we model the impact of the direction of upscaling or downscaling on productivity. The Average Treatment Effect on the Treated (ATT) is estimated using the following specification:

$$\log(Y_{it}) = \delta_1\, UPDIR_{it} + \delta_1\, DOWNDIR_{it} + \beta \log(TICK)_{it} + \alpha_i + \lambda_t + \epsilon_{it} \qquad (3.2)$$

The components $\alpha_i$ and $\lambda_t$ reflect the unobserved store fixed effects and date x day dummies to account for seasonal variation in the outcomes. The component $\varepsilon_{it}$ captures the idiosyncratic error that varies by store $i$ and time $t$. We include $\log(TICK)_{it}$, as a control in our model specification and cluster standard errors at the store x week level.

***Endogeneity issues.*** Real-time staffing adjustments are not randomly made and therefore exogenous. Given the autonomy to adjust real-time staffing levels, shift leads decide whether and how to adjust (up or down). Their decisions may be influenced by various factors that also affect the outcome variables (e.g. productivity). This self-selection bias introduces endogeneity, complicating the identification of the true causal effect of staffing adjustments. For instance, motivated or capable managers are more likely to adjust staffing effectively, directly impacting productivity. Proactive managers who anticipate demand fluctuations and adjust staffing accordingly can optimize productivity. These managerial characteristics influence both the decision to adjust labor hours and the resulting productivity, making the labor adjustments endogenous.

Simultaneity bias occurs when the dependent variable and independent variable mutually influence each other, creating a feedback loop. Managers adjust labor hours based on real-time performance metrics observed on a dashboard, such as sales, customer traffic, and service levels, to optimize store performance. These adjustments, in turn, directly impact the performance metrics they aim to optimize, such as improved customer service and increased sales. This bidirectional influence complicates the causal interpretation of the effect of labor adjustments on performance.

Omitted variable bias arises when our model excludes relevant variables that are correlated with both staffing adjustments and performance outcomes. Factors like managerial skill, local economic conditions, promotional activities, and store-specific characteristics can influence labor needs and performance. Their exclusion can lead to biased and inconsistent estimates.

To address endogeneity, we employ instrumental variables (IVs) that satisfy the relevance and exclusion restriction assumptions (Wooldridge, 2010). This IV approach allows us to obtain unbiased estimates of the effect of real-time schedule adjustments on store productivity while addressing potential sources of endogeneity. We need at least one IV for each endogenous variable: $UPDIR_{it}$ and $DOWNDIR_{it}$.

The relevance condition ensures that the instrument correlates with the endogenous regressors, $UPDIR_{it}$ and $DOWNDIR_{it}$, while the exclusion restriction requires that the instrument is uncorrelated with the error term, $\varepsilon_{it}$. This means the instrument can affect the outcome variable (e.g., productivity) only through the endogenous variable. By leveraging IVs, we can isolate the exogenous variation in labor adjustments, providing a clearer estimate of their causal effect on productivity and sales.

Assuming the validity of our instruments, we can obtain unbiased causal estimates using the standard two-stage least squares (2SLS) procedure (Wooldridge, 2010). However, due to the binary nature of our endogenous variables $UPDIR_{it}$ and $DOWNDIR_{it}$, the conditional expected function (CEF) associated with the first stage is likely nonlinear. Consequently, the usual linear first stage of the 2SLS procedure may not accurately approximate the true underlying CEF for these binary endogenous variables. Although the 2SLS estimators remain consistent, they may not be efficient. To enhance estimation efficiency, we employ the procedure outlined in Chapter 21 of Wooldridge (2010), commonly known as the Probit-2SLS procedure (Cerulli et al., 2014). This approach, also utilized by Kamalahmadi et al. (2021), involves three stages.

**Probit-Stage.** Run a Probit regression for the $UPDIR_{it}$ and $DOWNDIR_{it}$ variables on the instruments, $Z_{it} = (Instrument_{m,\,it})$ and the control variables, $X_{it}$

$$P(UPDIR_{it} = 1 \mid TREAT_{it} = 1) = \tag{3.3}$$
$$\phi\left(\beta_0^{up,\,0} + \beta_1^{up,\,0} Z_{it} + \beta_2^{up,\,0} \log(TICK_{it}) + \alpha_{it}^{up,\,0} + \lambda_{it}^{up,\,0} + \in_{it}^{up,\,0}\right)$$
$$P(DOWNDIR_{it} = 1 \mid TREAT_{it} = 1) =$$
$$\phi\left(\beta_0^{down,\,0} + \beta_1^{down,\,0} Z_{it} + \beta_2^{down,\,0} \log(TICK_{it}) + \alpha_{it}^{down,\,0} + \lambda_{it}^{down,\,0} + \in_{it}^{down,\,0}\right)$$

where $\phi$ is the cumulative distribution function (CDF) of the standard normal distribution. We obtain the corresponding predicted values, denoted as $\hat{P}(UPDIR_{it} = 1 \mid TREAT_{it} = 1)$ and $\hat{P}(DOWNDIR_{it} = 1 \mid TREAT_{it} = 1)$.

**2SLS-Stage 1.** Run an OLS regression for the

$$UPDIR_{it} = \beta_0^{up,\,1} + \beta_1^{up,\,1}\, \hat{P}(UPDIR_{it} = 1 \mid TREAT_{it} = 1) + \tag{3.4}$$
$$\beta_2^{up,\,1}\, \hat{P}(DOWNDIR_{it} = 1 \mid TREAT_{it} = 1) +$$
$$\beta_3^{up,\,1} \log(TICK_{it}) + \alpha_{it}^{up,\,1} + \lambda_{it}^{up,\,1} + \in_{it}^{up,1}$$
$$DOWNDIR_{it} = \beta_0^{down,\,1} + \beta_1^{down,\,1}\, \hat{P}(UPDIR_{it} = 1 \mid TREAT_{it} = 1) +$$

$\beta_2^{down,1}\ \hat{P}(DOWNDIR_{it} = 1 \mid TREAT_{it} = 1) +$

$\beta_3^{down,1}\log(TICK_{it}) + \alpha_{it}^{down,1} + \lambda_{it}^{down,1} + \epsilon_{it}^{down,1}$

We denote the corresponding predicted values as $\widehat{UPDIR}_{it}$ and $\widehat{DOWNDIR}_{it}$, respectively.

**2SLS-Stage 2.** We replace the $UPDIR_{it}$ and $DOWNDIR_{it}$ variables in Equation (1) by $\widehat{UPDIR}_{it}$ and $\widehat{DOWNDIR}_{it}$ variables and estimate it with OLS and robust standard errors that are clustered at week-store level.

$$\log(Y_{it}) = \beta_0 + \delta_1\ \widehat{UPDIR}_{it} + \delta_2\ \widehat{DOWNDIR}_{it} + \beta_3\ X_{it} + \epsilon_{it} \tag{3.5}$$

***Estimation with instrumental variables.*** We considered a large set of instrumental variables, whose definitions can be found in the appendix. Table 3.4 below displays the selected instrumental variables and their definitions:

These instruments are related to customer, sales, and labor-related metrics, and are derived using data from $t = t - 1$ (yesterday) or $t = t - 7$ (the same day last week). Additionally, we may use averages over the past seven days or the ratio of yesterday's values to the mean over the past seven days. It is common practice to use lagged variables as instruments (Mani et al. 2015; Perdikaki et al. 2012).

Customer metrics describe the level of traffic over the plan ($DevTICK_{it}$), average customer waiting time before being helped ($WAIT_{it}$), and average service time ($SERVICE_{it}$). For instance, relative deviation from the ticket forecast and average customer waiting time are directly tied to operational aspects that can influence labor adjustments. Peak service time ($MAX(SERVICE_{it})$) and the coefficient of variation of service times ($CV\_SERVICE_{it}$) indicate how managers might respond to changing conditions, ensuring that these instruments are correlated with $UPDIR_{it}$ and $DOWNDIR_{it}$.

Sales metrics, such as basket value ($BV_{it}$) and conversion rate ($CR_{it}$), directly relate to sales performance. These metrics can indicate the need for staffing adjustments to meet customer demand, providing relevant information for managerial decisions.

**Table 3.4** Instrument variables definitions

| | |
|---|---|
| **Customer related metrics:** | |
| Relative deviation from ticket forecast at t-1 and t-7 | $RelDevTICK_{it} = \frac{TICK_{it} - PlanTICK_{it}}{PlanTICK_{it}}$ |
| Average customer waiting time (minutes) at t-1 | $WAIT_{it} = \frac{1}{TICK_{it}} \sum_{j=1}^{TICK_{it}} wait_{ijt}$ where $wait_{ijt}$ is the customer wait time for ticket $j$ |
| Ratio of customer waiting time at t-1 to average for the past 7-days | $RATIO\_WAIT_{it} = \frac{WAIT_{i,t-1}}{\frac{1}{7}\sum_{t^*=t-7}^{t-1} WAIT_{it^*}}$ |
| Peak service time at t-1 | $MAX(SERVICE_{it})$ where $SERVICE_{it} = \frac{1}{SERV_{it}} \sum_{k=1}^{SERV_{it}} service_{ijt}$ and $service_{ijt}$ is the customer service time for ticket $j$ |
| Coefficient of variation of service times at t-1 | $CV\_SERVICE_{it}$ = ($SD\_SERVICE_{it}$ / $SERVICE_{it}$) x 100 where $SERVICE_{it} = \frac{1}{SERV_{it}} \sum_{k=1}^{SERV_{it}} service_{ijt}$ where $service_{ijt}$ is the customer service time for ticket $j$ and $SD\_SERVICE_{it}$ is the standard deviation of customer service times. |
| **Sales related metrics:** | |
| Basket value at t-1 and t-7 | $BV_{it} = SALES_{it} / TRANS_{it}$ |
| Average conversion rate for the past 7-days | $MEAN_CR_{it}$_l1_to_l7 = $\frac{1}{7}\sum_{t^*=t-7}^{t-1} \frac{TRANS_{it^*}}{TICK_{it^*}}$ |
| **Labor related metrics:** | |
| The number of floor staff and store management head count at t-1 | $EMP\_FLOOR\_MGMT_{i,t-1}$ |
| Ratio of store head count at t-1 to average for the past 7-days | $RATIO_EMP_{it}$_l1_to_l7 = $\frac{EMP_{i,\ t-1}}{\frac{1}{7}\sum_{t^*=t-7}^{t-1} EMP_{it^*}}$ |
| Deviation from planned labor hours at t-1 and t-7 | $DevLBR_{it} = LBR_{it} - PlanLBR_{it}$ |
| Relative deviation from planned workload (ticket per customer serving employee head count) at t-7 | $RelDevLOADS_{it} = \frac{DevLOADS_{it}}{PlanLOADS_{it}}$ where $DevLOADS_{it} = \frac{TICK_{it}}{SERVER_{it}} - \frac{PlanTICK_{it}}{PlanSERVER_{it}}$, and $PlanLOADS_{it} = \frac{PlanTICK_{it}}{PlanSERVER_{it}}$ |
| Employee idle rate per hour at t-1 | $IDLE_{it} =$ 1-($\sum_{k=1}^{SERV_{it}} service_{ikt}$ / $LBR\_floor_{it}$ )$*$ 100 where $service_{ikt}$ is the service time for ticket k and $SERV_{it}$ denotes the total number of tickets served |
| Ratio of employee idle rate per hour at t-1 to average for the past 7-days | $RATIO_IDLE_{it}$_l1_to_l7 = $\frac{IDLE_{i,\ t-1}}{\frac{1}{7}\sum_{t^*=t-7}^{t-1} IDLE_{it}}$ |
| Average productivity for the past 7-days | $MEAN_PROD_{it}$_l1_to_l7 = $\frac{1}{7}\sum_{t^*=t-7}^{t-1} PROD_{it}$ |
| Ratio of productivity at t-1 to average for the past 7-days | $RATIO_PROD_{it}$_l1_to_l7 = $\frac{PROD_{i,\ t-1}}{\frac{1}{7}\sum_{t^*=t-7}^{t-1} PROD_{it}}$ |

Labor-related metrics include the number of floor staff and store management headcount ($EMP\_FLOOR\_MGMT_{it}$), deviation from planned labor hours ($DevLBR_{it}$), and employee idle rate per hour ($IDLE_{it}$). These metrics provide insights into workforce management and operational efficiency, further influencing staffing decisions.

The relevance condition is satisfied as these instruments are directly related to operational and managerial decisions that influence staffing adjustments. By capturing aspects of customer interactions, sales performance, and labor management, these instruments ensure a robust correlation with $UPDIR_{it}$ and $DOWNDIR_{it}$. The F-statistics for the joint significance of the instruments in the first stage regressions indicate that the instruments for $UPDIR_{it}$ and $DOWNDIR_{it}$ are robust and not weak. Specifically, the F-statistic for $UPDIR_{it}$ is 28.86 and for $DOWNDIR_{it}$ is 19.75. Additionally, the under-identification test statistic is 30.93 and the weak identification test statistic is 27.10. These values, being well over the threshold of 10, suggest strong instrument relevance as per the criteria established by Staiger and Stock (1997).

To ensure these instruments satisfy the exclusion condition, they should not directly affect the outcome variable except through the endogenous variables. While average customer waiting time might influence manager decisions on labor adjustments, it should not directly impact productivity metrics independently of these decisions. Similarly, the basket value and conversion rate, although related to sales performance, influence labor adjustments without directly affecting productivity beyond these adjustments. The use of lagged variables, such as values from $t = t - 1$ or $t = t - 7$, is crucial as these past metrics cannot influence labor adjustments at time $t$ directly. This ensures that the instruments remain uncorrelated with the error term $\epsilon_{it}$.

By ensuring our instruments satisfy both the relevance and exclusion conditions, we can confidently use them to obtain unbiased estimates of the effects of real-time schedule adjustments on store productivity. This robust approach helps isolate the exogenous variation in labor adjustments, leading to clearer insights into their causal effects on productivity and sales.

***Estimation results.*** We present the estimation results of the model specified by Equation (3) in Table 3.5. Column 1-3 show the results of the OLS estimation. There is no significant association between increasing or decreasing real-time staffing levels and store productivity. Specifically, the coefficients for $UPDIR_{it}$ and $DOWNDIR_{it}$ in the OLS estimations are not statistically significant across productivity ($\log(PROD)$)), labor hours ($\log(LBR)$), and sales ($\log(SALES)$). This indicates that changes in real-time staffing levels do not have a discernible impact on these outcomes when using OLS estimation. However, the variable $\log(TICK)$ is consistently significant and positive across all three columns, suggesting that ticket volume is a strong predictor of productivity, labor hours, and sales.

The results from the OLS estimations suggest that simple increases or decreases in staffing levels, without considering endogeneity, do not significantly impact productivity, labor hours, or sales. This aligns with our expectation that the effectiveness of staffing adjustments requires a more sophisticated analysis to uncover true causal relationships.

As we discussed in Section 3.1.2, decisions to increase or decrease real-time staffing are endogenous. Thus, we report the causal estimates using the Probit-2SLS procedure in Columns 4-6. The results from this estimation provide a more detailed picture.

The coefficient for $UPDIR_{it}$ in Column 4 ($\log(PROD)$) is positive and statistically significant at the 5% level (0.162**). This indicates that increasing real-time staffing levels has a positive causal impact on store productivity, supporting Hypothesis 1 from our conceptual framework. Upscaling labor hours during high visitor traffic boosts productivity by increasing sales more than the labor hours. This positive impact on productivity highlights the effectiveness route, where additional labor during peak times enhances customer service and operational efficiency. Specifically, the expected store productivity uplift is 17.45% (95% CI [8.04%, 26.86%]).

In Column 6 ($\log(SALES)$), the coefficient for $UPDIR_{it}$ is positive but not statistically significant. In Column 5 ($\log(LBR)$), the coefficient for $UPDIR_{it}$ is negative and statistically significant coefficient at the 10% level (-0.106*), suggesting that increasing staffing levels may reduce labor hours, potentially indicating more efficient labor use. This finding indicates that stores were initially planned with fewer labor hours to match the actual traffic. Despite the upscaling of hours, stores were not able to fully close the labor hour gap, suggesting more efficient use of labor. The expected labor hours uplift is -10.30% (90% CI [-22.48 %, 1.87%]).

The positive sign for downscaling (though not statistically significant) in the same column suggests that reducing staffing levels may increase labor hours. This could occur because reduced staffing leads to longer working hours for the remaining employees to cover the shortfall, highlighting the trade-off between staffing adjustments and labor efficiency. These opposite signs for upscaling and downscaling are consistent with the explanation that optimizing staffing levels leads to more efficient labor use, while inadequate staffing may necessitate increased labor hours to maintain service levels.

**Table 3.5** Average treatment effect on the treated results of the model in (3)

| | OLS | | | Probit-2SLS | | |
|---|---|---|---|---|---|---|
| Variables | log(PROD) (1) | log(LBR) (2) | log(SALES) (3) | log(PROD) (4) | log(LBR) (5) | log(SALES) (6) |
| $UPDIR$ | 0.038 | -0.030 | 0.007 | 0.162** | -0.106* | 0.056 |
| | (0.031) | (0.020) | (0.020) | (0.048) | (0.074) | (0.471) |
| $DOWNDIR$ | -0.005 | 0.003 | -0.002 | 0.042 | 0.059 | 0.101 |
| | (0.018) | (0.011) | (0.014) | (0.662) | (0.324) | (0.206) |
| log($TICK$) | 0.656**** | 0.193*** | 0.849**** | 0.653*** | 0.246*** | 0.899*** |
| | (0.081) | (0.062) | (0.065) | (0.000) | (0.000) | (0.000) |
| Store fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Date x day of the week fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 607 | 607 | 607 | 607 | 607 | 607 |
| R-squared | 0.724 | 0.918 | 0.902 | 0.713 | 0.910 | 0.895 |
| Adjusted R-squared | 0.688 | 0.907 | 0.890 | 0.676 | 0.898 | 0.881 |
| | | | | F-statistics for $UPDIR$ | | 28.86 |
| | | | | F-statistics for $DOWNDIR$ | | 19.75 |
| | | | | Under identification test statistics | | 30.93 |
| | | | | Weak identification test statistics | | 27.10 |

*Notes.*
Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Conversely, the coefficients for $DOWNDIR_{it}$ in Columns 4-6 are not statistically significant, indicating that decreasing real-time staffing levels does not have a significant causal impact on productivity, labor hours, or sales when using the Probit-2SLS procedure. This partially supports Hypothesis 2, suggesting that downscaling labor hours during low traffic periods does not significantly harm productivity or sales. This aligns with the efficiency route but indicates that the primary gains are through upscaling during peak times.

We also note that the OLS results in Columns 1-3 of Table 3.5 underestimate the impact of real-time staffing adjustments. The Probit-2SLS estimates show that increasing staffing levels can significantly enhance productivity, which is not captured by the OLS estimates. This demonstrates the importance of addressing endogeneity to obtain unbiased and accurate estimates of the effects of staffing adjustments. The significant positive coefficient for log($TICK$) across both estimation methods further highlights the critical role of ticket volume in determining store performance metrics.

**Table 3.6** Average treatment effect on the treated results of the model in (3) excluding the post-experiment period

| | OLS | | | Probit-2SLS | | |
|---|---|---|---|---|---|---|
| Variables | log(PROD) (1) | log(LBR) (2) | log(SALES) (3) | log(PROD) (4) | log(LBR) (5) | log(SALES) (6) |
| $UPDIR$ | 0.046 | -0.029 | 0.017 | 0.214** | -0.166 ** | 0.048 |
| | (0.037) | (0.025) | (0.024) | (0.020) | (0.017) | (0.565) |
| $DOWNDIR$ | -0.004 | -0.002 | -0.007 | 0.027 | -0.017 | 0.010 |
| | (0.022) | (0.015) | (0.018) | (0.781) | (0.793) | (0.896) |
| log($TICK$) | 0.641 **** | 0.152 *** | 0.794 **** | 0.605*** | 0.196 *** | 0.800 *** |
| | (0.094) | (0.068) | (0.078) | (0.000) | (0.012) | (0.000) |
| Store fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Date x day of the week fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 439 | 439 | 439 | 439 | 439 | 439 |
| R-squared | 0.735 | 0.910 | 0.904 | 0.676 | 0.887 | 0.890 |
| Adjusted R-squared | 0.697 | 0.897 | 0.890 | 0.717 | 0.901 | 0.904 |
| | | | | F-statistics for $UPDIR$ | | 21.31 |
| | | | | F-statistics for $DOWNDIR$ | | 16.38 |
| | | | | Under identification test statistics | | 26.24 |
| | | | | Weak identification test statistics | | 22.58 |

*Notes.*
Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

We conducted a robustness check by focusing solely on the experiment period and excluding the post-experiment period. The results are presented in Table 3.6. In comparing the estimates with those from the previous analysis including the post-experiment period, we find similar but stronger results in the robustness checks. The OLS estimates displayed in columns 1-3 show slightly higher coefficients for $UPDIR_{it}$ and $DOWNDIR_{it}$ but these remain statistically insignificant. The Probit 2-SLS estimates in columns 4-6 indicate a stronger positive impact on labor hours, while the effect on sales remains insignificant. $UPDIR_{it}$ consistently shows a significant positive impact on productivity and a significant negative impact on labor hours, while its effect on sales remains insignificant in both analyses. This reinforces the robustness of our findings and the effectiveness of the Probit-2SLS approach in providing reliable causal estimates while addressing endogeneity.

## 3.7 Discussion

This study provides significant insights into the impact of real-time staffing adjustments on retail productivity. Allowing stores to adjust staffing levels in real-time showed a notable 6.24% increase in productivity. Compared to other retail trade industries in 2022—non-store retailers saw an 11.1% productivity change, electronics and appliance stores 2.5%, and health and personal care stores only 0.3% according to U.S. Bureau of Labor Statistics—this increase highlights the substantial potential of real-time staffing adjustments. Practical benefits include better alignment of labor supply with actual customer demand, more efficient use of labor hours, and reduced overstaffing and understaffing.

To implement real-time staffing adjustments effectively, managers should train shift leads to make informed decisions based on real-time data. Tools like dashboards providing near real-time metrics on store traffic, sales, and service levels can equip shift leads with the necessary information to optimize staffing levels dynamically. However, balancing flexibility with the desire for predictable schedules is crucial. Managers must communicate clearly with employees about the benefits of real-time adjustments and address any concerns about potential disruptions. Resistance from both employees and managers accustomed to traditional scheduling practices can be mitigated with adequate training and support.

Real-time staffing adjustments can benefit various retail sectors. For example, apparel stores can handle peak shopping times during seasonal sales such as Christmas and Black Friday. Grocery stores can manage high-traffic periods like weekends and holidays, and department stores can optimize staffing during clearance events. The effectiveness of real-time staffing adjustments may vary between urban and rural areas due to differences in customer traffic patterns. Urban stores with higher and more unpredictable traffic may see more significant benefits compared to rural stores with more stable traffic. Additionally, this approach can be effectively implemented within the strict labor laws and regulations of the country, ensuring that flexible staffing strategies comply with legal requirements while optimizing operational efficiency.

Real-time staffing adjustments can have long-term impacts on employee satisfaction and retention. They can increase job satisfaction by providing employees with the flexibility to adjust work hours based on personal needs, enhance engagement and performance through greater autonomy, and foster a supportive work environment through training and clear guidelines. However, challenges include potential stress and uncertainty, which can be mitigated by clear communication and fair practices, ensuring fairness through transparent criteria for selecting employees for additional hours or early leave.

Future research should broaden the scope of this study by applying the real-time staffing adjustment framework to other retail sectors such as apparel, groceries, electronics, and department stores. These sectors experience significant fluctuations in customer demand and can benefit from dynamic staffing adjustments. For instance, apparel retail can examine peak shopping times during seasonal sales, while grocery stores can focus on high-traffic periods like weekends and holidays. Electronics and technology stores can analyze customer service needs during product launches, and department stores can assess staffing needs during clearance events.

Measuring employee satisfaction and its impact on productivity is crucial for a comprehensive understanding. Future studies should employ various methods to measure employee satisfaction more effectively. Regular, anonymous surveys and questionnaires can provide insights into overall job satisfaction and work environment perceptions. In-depth interviews and focus groups can offer a deeper understanding of specific issues and areas for improvement. Monitoring turnover rates and tracking absenteeism and tardiness can serve as indirect indicators of employee dissatisfaction or burnout. Pulse surveys can capture real-time feedback on specific aspects of their work environment and changes, providing immediate data to inform staffing adjustments.

Incorporating customer feedback and satisfaction metrics would offer a more comprehensive view of the impact on overall store performance. Key metrics include the Net Promoter Score (NPS) for measuring customer loyalty, Customer Effort Score (CES) for ease of interaction, and Customer Satisfaction (CSAT) surveys for evaluating specific purchase or service contentment. Mystery shopping can offer unbiased insights, while analyzing customer complaints and feedback can identify recurring issues. Additionally, wait time analysis during peak periods can assess the effectiveness of staffing adjustments in maintaining service levels.

To assess the sustainability of productivity gains, future research should conduct longitudinal studies to determine if productivity gains are maintained over time. Repeat experiments in different contexts or seasons can validate the consistency of productivity gains. Comparative analysis of stores that implemented real-time adjustments with those that did not over an extended period can highlight long-term impacts. Continuous monitoring of key performance indicators (KPIs) such as sales per hour, customer satisfaction, and employee engagement will provide ongoing insights. Additionally, assessing long-term effects on employee well-being, including job satisfaction and stress levels, and measuring long-term changes in operational efficiency, such as reduced overtime costs and improved service levels, will demonstrate the sustainability of productivity improvements.

Future research should focus on collecting more granular data on employee activities. Implementing systems that track the specific tasks employees perform during their shifts, such as mobile apps for logging activities or wearable technology, can provide a more detailed understanding of labor allocation and the effectiveness of shift leads. Regular audits and gathering employee feedback can help verify the alignment between planned and actual activities, offering deeper insights into operational efficiency and the impact of real-time staffing adjustments. Enhancing data collection methods and obtaining a more detailed breakdown of roster activities can better understand labor allocation dynamics and the critical role of shift leads in optimizing store productivity.

One limitation is the lack of incentives for maximizing sales per hour worked through intraday staffing adjustments. Formal contracts that include specific incentives for maximizing sales per hour worked can significantly increase the motivation and effort of shift leads and employees. These contracts can outline clear performance metrics and rewards, such as bonuses or other financial incentives, for meeting or exceeding sales targets. This structured approach fosters a more driven and goal-oriented work environment. In the absence of formal contracts with performance-based incentives, employees may lack the motivation to go beyond their basic job requirements.

At the end of the experiment, we could not collect feedback from employees regarding the real-time staffing adjustments. Understanding employees' perspectives on being offered the opportunity to extend shifts or leave early is crucial. If productivity increases but workers are unhappy, it could lead to higher attrition rates. While shift leads make decisions to adjust hours and seek volunteers, employees are free to accept or decline these requests. The pressure potentially experienced to accommodate a supervisor's request may negatively impact employee job satisfaction. Conversely, the flexibility offered to employees, who have the right to accept or decline the request, may help them meet personal needs, such as earning extra money or having time to pick up children from daycare. Unfortunately, our retail partner was not open to conducting employee surveys, preventing us from quantifying the impact of intraday hour adjustments on employee satisfaction.

To address this limitation, we plan to conduct interviews with employees to gather qualitative feedback on their experiences with real-time staffing adjustments. Additionally, we will consider evaluating indirect indicators of employee satisfaction, such as absenteeism and turnover rates. To further enrich our understanding, we may also analyze Google reviews and Net Promoter Score (NPS) data to capture broader customer and employee sentiments. These methods will provide a more comprehensive understanding of how staffing adjustments impact employee well-being and overall job satisfaction.

Adjusting hours may affect customer satisfaction scores of a store. However, our ability to measure this causal relationship is limited due to the low volume of customer surveys during the six-week period. This limitation hinders our capacity to fully understand the effect of real-time staffing adjustments on the customer experience.

There are also interesting considerations regarding the extent to which we would find similar effects in sectors. In labor-intensive industries, reducing working hours has been shown to improve efficiency, whereas in knowledge-intensive industries, it may decrease efficiency (Kim et al., 2022). This suggests that the impact of reduced labor hours varies significantly by industry type. Further research in knowledge-intensive industries, such as financial consulting, and insurance sectors requiring specialized knowledge and analysis, could provide more insights.

In summary, our study highlights the potential of real-time staffing adjustments to improve store productivity and operational efficiency. Adaptive labor management strategies are essential for modern retail operations, where fluctuating customer demand necessitates a flexible approach to staffing. We encourage further exploration and application of real-time staffing adjustments to optimize performance and operational efficiency. The insights gained from this research can guide retail managers in implementing dynamic staffing strategies that balance operational efficiency with effective labor management.

# CONCLUSION

The exploration of omnichannel retailing, real-time staffing adjustments, and the impact of physical store openings on return rates yields significant insights for the retail industry. Through our comprehensive investigation, we have illuminated how the strategic integration of physical and online channels and innovative labor management practices can collectively enhance operational efficiency.

Firstly, our research into the value of experience-centric stores in omnichannel retail highlights the importance of physical presence in driving revenue growth and improving customer experience. For online-first retailers, particularly those dealing with multi-brand assortments, the expansion into large, experience-centric stores has shown substantial benefits. These stores facilitate sensory product inspections, provide access to expert advice, offer instant gratification, and streamline returns. Such benefits are crucial in differentiating the in-store experience from online shopping, thereby fostering stronger customer engagement and loyalty. The empirical evidence from our study shows that while short-term revenue uplifts are influenced by perceived utilities and category-specific intentions, long-term uplifts are shaped by a deeper integration of store offerings with customer expectations and behaviors. These findings provide valuable guidance for retailers in planning store formats and assortments to optimize cross-channel synergies and maximize total sales.

Secondly, our investigation into the effects of omnichannel store openings on return rates reveals the interplay between physical store presence and return behaviors. Our findings indicate that the introduction of physical stores has significantly impacted returns, with a noticeable shift in consumer behavior towards more frequent returns. Specifically, return rates rose by 28.0% for Large Store 1 and 21.7% for Large Store 2. This substantial increase suggests that the enhanced convenience of the return process is influencing the retailer's operational strategy. The analysis further reveals that the uplift in higher average price of returned items compared to sold items amplifies the effect of increased return rates, resulting in a more pronounced uplift in the return fraction.

Consequently, net revenue grows less than gross revenue, highlighting the delicate balance between increasing revenue and controlling return rates. To manage return rates effectively and improve net revenue, retailers should implement strategies such as enhanced product descriptions, robust post-purchase support, and targeted approaches for high-value items. These measures not only help reduce return rates but also enhance customer satisfaction, which is critical for long-term sustainability.

Finally, the study on real-time staffing adjustments demonstrates the potential of adaptive labor management to boost store productivity. The ability to align staffing levels with real-time customer demand significantly enhances operational efficiency, reducing both overstaffing and understaffing. The implementation of real-time staffing adjustments requires careful consideration of employee well-being, clear communication, and adequate training. Our findings indicate that such adjustments lead to a notable increase in productivity without negatively impacting sales volume.

Interestingly, the productivity boost is mostly observed during upscaling, even though upscaling is done less frequently than downscaling. This discrepancy likely arises because upscaling is seen as riskier and harder to justify to headquarters. Additionally, it is more challenging to implement, as extending labor hours beyond a certain point can be difficult. For example, a 5-hour upscaling may be easier and more acceptable for employees to accommodate voluntarily compared to a 10-hour extension. This highlights the importance of strategic planning and employee buy-in when implementing flexible and responsive labor management strategies in modern retail operations, where customer demand fluctuates frequently.

In conclusion, the insights from our research offer a robust set of findings on how and when to best open stores, efficiently manage personnel, and curb the increase in return rates due to increased return convenience. By leveraging the complementarities of online and offline channels and adopting innovative labor management practices, retailers can enhance customer satisfaction, improve operational efficiency, and drive sustainable revenue growth. Future research should continue to explore these dimensions, particularly focusing on the long-term impacts of these strategies across different retail sectors and geographic regions. The ongoing evolution of consumer behavior and technological advancements will undoubtedly present new challenges and opportunities for omnichannel retailing, making continuous adaptation and innovation essential for success.

# BIBLIOGRAPHY

Anderson E, Hansen K, Simester D (2009) The option value of returns: Theory and empirical evidence. *Marketing Science* 28(3):405–423.

Akturk MS, Ketzenberg M, Heim GR (2018) Assessing impacts of introducing ship-to-store service on sales and returns in omnichannel retailing: A data analytics study. *Journal of Operations Management* 61(1):15–45.

Angrist JD, Pischke JS (2008) Mostly Harmless Econometrics: An Empiricist's Companion, 1st ed. (Princeton University Press, Princeton, NJ).

Ansari A, Mela CF, Neslin SA (2008) Customer Channel Migration. *Journal of Marketing Research* 45(1):60–76.

Athey S, Imbens G (2022) Design-Based Analysis in Difference-in-Differences Settings with Staggered Adoption. *Journal of Econometrics* 226(1):62-79.

Avery J, Steenburgh TJ, Deighton J, Caravella M (2012) Adding Bricks to Clicks: Predicting the Patterns of Cross-Channel Elasticities over Time. *Journal of Marketing* 76(3):96–111.

Balakrishnan A, Sundaresan S, Zhang B (2014) Browse-and-Switch: Retail-Online Competition under Value Uncertainty. *Production and Operations Management* 23(7):1129–1145.

Basker E (2005) Job creation or destruction? Labor market effects of Wal-Mart expansion. *The Review of Economics and Statistics* 87(1):174–183.

Bell DR, Gallino S, Moreno A (2014) How to Win in an Omnichannel World. *MIT Sloan Management Review*:11.

Bell DR, Gallino S, Moreno A (2018) Offline Showrooms in Omnichannel Retail: Demand and Operational Benefits. *Management Science* 64(4):1629–1651.

Bell DR, Gallino S, Moreno A (2020) Customer Supercharging in Experience-Centric Channels. *Management Science* 66(9):4096–4107.

Bernon M, Cullen J, Gorst J (2016) Online retail returns management. *International Journal of Physical Distribution & Logistics Management* 46(6/7):584–605.

Biron B (2020) Ikea is opening 50 small-format stores globally as it experiments with new retail concepts in “vibrant urban destinations” during the pandemic. Business Insider. Retrieved (October 22, 2022), https://www.businessinsider.com/ikea-to-open-new-stores-urban-areas-cities-amid-pandemic-2020-10.

Biyalogorsky E, Naik P (2003) Clicks and Mortar: The Effect of On-line Activities on Off-line Sales. *Marketing Letters*:12.

Callaway B, Sant’Anna PHC (2021) Difference-in-Differences with multiple time periods. *Journal of Econometrics* 225(2):200–230.

Cerulli G (2014) ivtreatreg: A command for fitting binary treatment models with heterogeneous response to treatment and unobservable selection. Stata J. 14(3):453–480.

Chan CW, Farias VF, Escobar GJ (2017) The impact of delays on service times in the intensive care unit. *Management Science* 63(7):2049–2072.

Chapados N, Joliveau M, L’Ecuyer P, Rousseau LM (2014) Retail store scheduling for profit. *European Journal of Operational Research* 239(3):609–624.

Chuang HHC, Oliva R, Perdikaki O (2016) Traffic-based labor planning in retail stores. *Production and Operations Management* 25(1):96–113.

de Chaisemartin C, D’Haultfœuille X (2020) Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review* 110(9):2964–2996.

Kim Do-Hoe, Jesun Yeon (2022) Different effects of working hour reduction on labor-intensive and knowledge-intensive industries in the era of artificial intelligence: a meta-frontier approach. *Applied Economics* 55: 2493 - 2504.

Edwards B (2021) A tale of two Apple Stores (the first two). *Macworld*. Retrieved (November 10, 2022), https://www.macworld.com/article/212229/first2apple-stores.html.

Ertekin N (2018) Immediate and Long-Term Benefits of In-Store Return Experience. *Production and Operations Management* 27(1):121–142.

Easton FF, Goodale JC (2005) Schedule recovery: Unplanned absences in service operations. *Decision Sciences* 36(3):459–488.

Fishbein, M., & Ajzen, I. (1975). Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, MA: Addison-Wesley.

Fisher M, Krishnan J, Netessine S (2006) Retail store execution: An empirical study. Preprint, submitted December 3, and posted on September 4, 2013.

Frasquet M, Miquel-Romero MJ (2021) Competitive (versus loyal) showrooming: An application of the push-pull-mooring framework. *Journal of Retailing and Consumer Services* 62:102639.

Gallino S, Moreno A (2014) Integration of Online and Offline Channels in Retail: The Impact of Sharing Reliable Inventory Availability Information. *Management Science* 60(6):1434–1451.

Gans N, Koole G, Mandelbaum A (2003) Telephone call centers: a tutorial and literature review. *Manufacturing and Service Operations Management* 5(2): 79-141.

Gao F, Su X (2016) Online and offline information for omnichannel retailing. *SSRN Electronic Journal*.

Gao F, Su X (2017) Omnichannel retail operations with buy-online-and-pick-up-in-store. *Management Science* 63(8):2478–2492.

Geyskens I, Gielens K, Dekimpe MG (2002) The Market Valuation of Internet Channel Additions. *Journal of Marketing* 66(2):102–119.

Goodman-Bacon A (2021) Difference-in-differences with variation in treatment timing. *Journal of Econometrics* 225(2):254–277.

Henly JR, Lambert SJ (2014) Unpredictable work timing in retail jobs: Implications for employee work–life conflict. *ILR Review* 67(3):986–1016.

Herrera S (2022) Amazon Set to Close Bookstores, Other Shops in Retail Shift to Groceries, Fashion. *WSJ*. Retrieved (September 28, 2022), https://www.wsj.com/articles/amazon-set-to-close-bookstores-other-shops-in-retail-shift-to-groceries-fashion-11646252544.

Hirche C, Bijmolt T, Gijsenberg M (2022) When offline stores reduce online returns. *Sustainability* 14(13):7829.

Hur D, Mabert VA, Bretthauer KM (2004) Real-time work schedule adjustment decisions: An investigation and evaluation. *Production and Operations Management* 13(4):322–339.

Iravani, S. M., M. P. Van Oyen, and K. T. Sims (2005). Structural flexibility: A new perspective on the design of manufacturing and service operations. *Management Science* 51(2), 151-166.

Kamalahmadi M, Yu Q, Zhou YP (2021) Call to Duty: Just-in-time scheduling in a restaurant chain. *Management Science* 67(11):6751–6781.

Kapner S (2023) Retailers Clamp Down on Returns. *WSJ*. Retrieved (February 20, 2024), https://www.wsj.com/articles/online-retailers-tighten-return-policies-to-boost-profits-9bf6ccc2.

Kapner S (2021) E-Commerce Needs Real Store Locations Now More Than Ever. *WSJ*. Retrieved (August 15, 2022), https://www.wsj.com/articles/e-commerce-needs-real-store-locations-now-more-than-ever-11637836200.

Kesavan S, Lambert SJ, Williams JC, Pendem PK (2022) Doing well by doing good: Improving retail store performance with responsible scheduling practices at the Gap, Inc. *Management Science* 68(11):7818–7836.

Kesavan S, Staats BR, Gilland W (2014) Volume flexibility in services: The costs and benefits of flexible labor resources. *Management Science* 60(8):1884–1906.

Kesavan S, Mani V (2015) An overview of industry practice and empirical research in retail workforce management. Agrawal N, Smith SA, eds. Retail Supply Chain Management. (Springer, Boston), 113–145.

King K (2022) Macy's Is Betting Even Bigger on Smaller Stores. WSJ. Retrieved (August 26, 2022), https://www.wsj.com/articles/macys-is-betting-even-bigger-on-smaller-stores-11651579200.

Kong X, Zhou S, Liu C (2017) The impact of return policies on customer loyalty and the moderating effect of return experience. *Journal of Retailing and Consumer Services* 38:1–10.

Krischer MM, Penney LM, Hunter EM (2010) Can counterproductive work behaviors be productive? cwb as emotion-focused coping. J. Occupational Health Psych. 15(2):154.

Kumar A, Mehra A, Kumar S (2019) Why Do Stores Drive Online Sales? Evidence of Underlying Mechanisms from a Multichannel Retailer. *Information Systems Research* 30(1):319–338.

Lambert SJ, Fugiel PJ, Henly JR (2014) Precarious work schedules among early-career employees in the US: A national snapshot. EINet research brief, University of Chicago, Chicago.

Lee C, Morewedge C (2023) Mental accounting of product returns. *Journal of Consumer Psychology* 33(3):583–590.

Lee HS (Huck), Kesavan S, Deshpande V (2021) Managing the impact of fitting room traffic on retail sales: Using labor to reduce phantom stockouts. *Manufacturing and Service Operations Management* 23(6):1580–1596.

Lee J (2022) How Bricks Might Save Clicks. *WSJ*. Retrieved (August 23, 2022), https://www.wsj.com/articles/how-bricks-might-save-clicks-11650619815.

Lee J (2023) It’s Not Your Imagination—Shopping in Person Is Getting Worse. *WSJ*. Retrieved (September 14, 2023), https://www.wsj.com/business/retail/theft-shrink-shopping-locked-up-products-bb40ec70.

Li G, Pan X (2023) Optimal return shipping insurance policy with consumers' anticipated regret. *Production and Operations Management* 32(10):3209–3226.

Li M, Choudhury A (2020) Using website information to reduce postpurchase dissonance: A mediated moderating role of perceived risk. *Psychology and Marketing* 38(1):56–69.

Li M, Liu Y (2021) Beneficial product returns in supply chains. *Production and Operations Management* 30(11):3849–3855.

Li Y (2023) Research online and purchase offline: The disruptive impact of consumers' online information on offline sales interaction. *Psychology and Marketi*ng 40(12):2642–2652.

Liu J, Xu Q (2020) Joint decision on pricing and ordering for omnichannel BOPS retailers: Considering online returns. *Sustainability* 12(4):1539.

Locke EA, Latham GP (2002) Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist* 57(9):705–717.

Locke EA, Latham GP (2006) New directions in goal-setting theory. *Curr Dir Psychol Sci* 15(5):265–268.

Lu G, Du RY, Peng X (David) (2022) The impact of schedule consistency on shift worker productivity: An empirical investigation. *Manufacturing and Service Operations Management* 24(5):2780–2796.

Mac-Vicar M, Ferrer JC, Muñoz JC, Henao CA (2017) Real-time recovering strategies on personnel scheduling in the retail industry. *Computers & Industrial Engineering* 113:589–601.

Mani V, Kesavan S, Swaminathan JM (2015) Estimating the impact of understaffing on sales and profitability in retail stores. *Production and Operations Management* 24(2):201–218.

McKeever P (2022) U.S. retailers announced nearly seven times as many store openings as closings in the first quarter of 2022. *NRF*. Retrieved (August 17, 2022), https://nrf.com/blog/us-retailers-announced-nearly-seven-times-many-store-openings-closings-first-quarter-2022.

Mehrotra V, Ozlük O, Saltzman R (2010) Intelligent procedures for intra-day updating of call center agent schedules. *Production and Operations Management* 19(3):353–367.

Miranda L (2022) Online brands open more stores in the suburbs to be closer to customers. NBC News. Retrieved (August 17, 2022), https://www.nbcnews.com/business/online-brands-open-stores-suburbs-rcna23030.

Musalem A, Olivares M, Schilkrut A (2021) Retail in high definition: Monitoring customer assistance through video analytics. *Manufacturing and Service Operations Management* 23(5):1025–1042.

NRF (2023) 2023 Consumer Returns in the Retail Industry. Retrieved (February 20, 2024), https://nrf.com/research/2023-consumer-returns-retail-industry.

Netessine S, Fisher ML, Krishnan J (2010) Labor planning, execution, and retail store performance: An exploratory investigation. *Working Paper*:40.

Ofek E, Katona Z, Sarvary M (2011) “Bricks and Clicks”: The Impact of Product Returns on the Strategies of Multichannel Retailers. Marketing Science 30(1):42–60.

Ostrom A, Iacobucci D (1995) Consumer Trade-Offs and the Evaluation of Services. *Journal of Marketing* 59(1):13.

Parasuraman A, Zeithaml VA, Berry LL (1988) SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*.

Peng X, Ye Y, Ding X, Chandrasekaran A (2023) The impact of nurse staffing on turnover and quality: An empirical examination of nursing care within hospital units. *J of Ops Management* 69(7):1124–1152.

Perdikaki O, Kesavan S, Swaminathan JM (2012) Effect of traffic on sales and conversion rates of retail stores. *Manufacturing and Service Operations Management* 14(1):145–162.

Phillips EE (2017) Retailers Offer Myriad Returns Options to Retain Customers. Wall Street Journal (December 26) https://www.wsj.com/articles/retailers-offer-myriad-returns-options-to-retain-customers-1514308299.

Pozzi A (2013) The effect of Internet distribution on brick-and-mortar sales. *RAND Journal of Economics* 44(3):569–583.

Rooderkerk RP, Kök GA (2019), Omnichannel Assortment Planning. *Operations in an omnichannel world*. Eds S. Gallino and A. Moreno, 51-86.

Roth J, Sant’Anna PHC, Bilinski A, Poe J (2023) What’s trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics* 235(2):2218–2244.

Schneider D, Harknett K (2019) Consequences of routine work-schedule instability for worker health and well-being. *Am Sociol Rev* 84(1):82–114.

Shang G, Pekgün P, Ferguson M, Galbreth M (2017) How much do online consumers really value free product returns? Evidence from eBay. *Journal of Operations Management* 53-56(1):45–62.

Shoulberg W (2021) Wayfair Will Start to Open Stores in 2022...Finally. *Forbes.* Retrieved (October 22, 2022), https://www.forbes.com/sites/warrenshoulberg/2021/12/07/wayfair-will-start-to-open-stores-in-2022finally/.

Staiger DO, Stock JH (1997) Instrumental variables regression with weak instruments. Econometrica 65(3):557–586.

Sun L, Abraham S (2021) Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225(2):175–199.

Tan TF, Netessine S (2014) When does the devil make work? An empirical study of the impact of workload on worker productivity. *Management Science* 60(6):1574–1593.

Thompson, G., M. (1999). Labor scheduling, Part 4. *Cornell Hotel and Restaurant Administration Quarterly* 40(3): 85-96.

Ton Z (2009) The effect of labor on profitability: The role of quality. *Working Paper* Harvard Business School, Cambridge, MA.

Ton Z, Huckman RS (2008) Managing the impact of employee turnover on performance: The role of process conformance. *Organization Science* 19(1):56–68.

Wang K, Goldfarb A (2017) Can Offline Stores Drive Online Sales? *Journal of Marketing Research* 54(5):706–719.

Wolfe R (2022) Good Luck Returning Your Unwanted Clothes and Electronics. Wall Street Journal (November 1) https://www.wsj.com/articles/online-shopping-gap-zara-returns-exchanges-11667240388.

Womack JP, Jones DT (1997) Lean thinking—Banish waste and create wealth in your corporation. *J Oper Res Soc* 48(11):1148–1148.

Wooldridge JM (2010) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA).

Xia Y, Xiao T, Zhang G (2016) The impact of product returns and retailer's service investment on manufacturer's channel strategies. *Decision Sciences* 48(5):918–955.

Yoon S, Jeong J (2021) The effect of return policy on customers' purchase decisions and return behavior: Evidence from a natural experiment. *Journal of Business Research* 124:27–35.

Zhang DJ, Dai H, Dong L, Wu Q, Guo L, Liu X (2019) The Value of Pop-Up Stores on Retailing Platforms: Evidence from a Field Experiment with Alibaba. *Management Science* 65(11):5142–5151.

# Appendix B

**Table B1** Definitions of variables used in derivation of equations

| **Variable** | **Definition** |
|---|---|
| $GR_{before}$ | the gross revenues before store opening |
| $GR_{after}$ | the gross revenues after store opening |
| $RR_{before}$ | the return revenues before store opening |
| $RR_{after}$ | the return revenues after store opening |
| $NR_{before}$ | the net revenues before store opening $= GR_{before} - RR_{before}$ |
| $NR_{after}$ | the net revenues after store opening $= GR_{after} - RR_{after}$ |
| $RF_{before}$ | fraction of gross revenues returned (return fraction) before store opening $= \frac{RR_{before}}{GR_{before}}$ |
| $RF_{after}$ | fraction of gross revenues returned (return fraction) after store opening $= \frac{RR_{after}}{GR_{after}}$ |
| $RRate_{before}$ | fraction of items returned before store opening $= \frac{UR_{before}}{US_{before}}$ |
| $RRate_{after}$ | fraction of items returned after store opening $= \frac{UR_{after}}{US_{after}}$ |
| $Uplift_{GR}$ | the uplift of gross revenues due to store opening |
| $Uplift_{NR}$ | the uplift of net revenues due to store opening |
| $Uplift_{RR}$ | the uplift of return revenues due to store opening |
| $Uplift_{RF}$ | the uplift of return fraction due to store opening |
| $Uplift_{RRate}$ | the uplift of return rate due to store opening |

## B.1 Derivation of Equation (2.1)

This section provides a mathematical derivation of Equation (2.1), which quantifies the discrepancy between net revenue uplift and gross revenue uplift as a result of opening new stores or changes in sales strategies.

$$Uplift_{NR} = \frac{NR_{after} - NR_{before}}{NR_{before}} = \frac{\left(GR_{after} - RR_{after}\right) - \left(GR_{before} - RR_{before}\right)}{GR_{before} - RR_{before}}$$

$$= \frac{GR_{after} - GR_{before}}{GR_{before} - RR_{before}} - \frac{RR_{after} - RR_{before}}{GR_{before} - RR_{before}}$$

$$= \frac{\frac{\left(GR_{after} - GR_{before}\right)}{GR_{before}}}{\frac{\left(GR_{before} - RR_{before}\right)}{GR_{before}}} - \frac{\frac{\left(RR_{after} - RR_{before}\right)}{RR_{before}}}{\frac{\left(GR_{before} - RR_{before}\right)}{RR_{before}}}$$

$$= \frac{GR_{before}}{GR_{before} - RR_{before}} * Uplift_{GR} - \frac{RR_{before}}{GR_{before} - RR_{before}} * Uplift_{RR}$$

$$= \frac{GR_{before}}{GR_{before} - RR_{before}} * Uplift_{GR} - \frac{RR_{before}}{GR_{before} - RR_{before}} * Uplift_{GR}$$

$$- \frac{RR_{before}}{GR_{before} - RR_{before}} * \left(Uplift_{RR} - Uplift_{GR}\right)$$

$$= Uplift_{GR} - \frac{RR_{before}}{GR_{before} - RR_{before}} * \left(Uplift_{RR} - Uplift_{GR}\right)$$

$$= Uplift_{GR} - \frac{\frac{RR_{before}}{GR_{before}}}{\frac{\left(GR_{before} - RR_{before}\right)}{GR_{before}}} * \left(Uplift_{RR} - Uplift_{GR}\right)$$

$$= Uplift_{GR} - \frac{RF_{before}}{1 - RF_{before}} * \left(Uplift_{RR} - Uplift_{GR}\right) \tag{2.1}$$

## B.2 Derivation of Equation (2.2)

This section elaborates on the mathematical derivation of Equation (2.2), which represents the excess uplift of return revenues as influenced by the uplift in the return fraction.

$$Uplift_{RR} - Uplift_{GR} = \frac{RR_{after}}{RR_{before}} - 1 - \left(\frac{GR_{after}}{GR_{before}} - 1\right) = \frac{RR_{after}}{RR_{before}} - \frac{GR_{after}}{GR_{before}}$$

$$= \left(\frac{\frac{RR_{after}}{GR_{after}}}{\frac{RR_{before}}{GR_{before}}}\right) * \frac{GR_{after}}{GR_{before}} - \frac{GR_{after}}{GR_{before}}$$

$$= \left(\frac{RF_{after}}{RF_{before}} * \frac{GR_{after}}{GR_{before}}\right) - \frac{GR_{after}}{GR_{before}}$$

$$= \frac{GR_{after}}{GR_{before}} * \left(\frac{RF_{after}}{RF_{before}} - \frac{RF_{before}}{RF_{before}}\right) = \frac{GR_{after}}{GR_{before}} * (Uplift_{RF})$$

$$= \frac{GR_{after}}{GR_{before}} * Uplift_{RF}$$

$$= Uplift_{RF}\,(Uplift_{GR} + 1) \qquad (2.2)$$

## B.3 Derivation of Equation (2.3) and (2.4)

This section details the mathematical derivation of Equations (2.3) and (2.4), which further refines the discrepancy between net revenue and gross revenue uplifts by integrating the effect of the return fraction uplift. It substitutes the results of Equation (2.2) into Equation (2.1) to demonstrate how the return fraction uplift modifies the net revenue impact.

$$Uplift_{NR} = Uplift_{GR} - \left(\frac{RF_{before}}{1 - RF_{before}}\right) * \left(Uplift_{RF}\,(Uplift_{GR} + 1)\right)$$

$$= Uplift_{GR} - \left(\frac{RF_{before}}{1 - RF_{before}}\right) * Uplift_{RF} * Uplift_{GR}$$

$$- \left(\frac{RF_{before}}{1 - RF_{before}}\right) * Uplift_{RF}$$

$$= Uplift_{GR}\left[1-\left(\frac{RF_{before}}{1-RF_{before}}\right) * Uplift_{RF}\right]-\left(\frac{RF_{before}}{1-RF_{before}}\right) * Uplift_{RF}$$

$$Uplift_{NR} = [1 - C_{GR\to NR}]\, Uplift_{GR} - C_{GR\to NR} \tag{2.3}$$

$$\text{where } C_{GR\to NR} = \left(\frac{RF_{before}}{1-RF_{before}}\right) * Uplift_{RF} \tag{2.4}$$

**Table B2** Definitions of variables used in derivation of equations

| Variable | Definition |
|---|---|
| $US_{before}$ | the unit sales (=number of items sold) before store opening |
| $US_{after}$ | the unit sales (=number of items sold) after store opening |
| $UR_{before}$ | the number of items returned before store opening |
| $UR_{after}$ | the average price per item sold after store opening |
| $AS_{before}$ | the average price per item sold before store opening |
| $AS_{after}$ | the average price per item sold after store opening |
| $AR_{before}$ | the average price per returned item before store opening |
| $AR_{after}$ | the average price per returned item after store opening |

## B.4 Derivation of Equation (2.6)

This section explains the derivation of Equation (2.6), which decomposes the return fraction uplift in terms of the return rate uplift. This equation establishes a relationship between the changes in the rates of returns and the prices of sold and returned items, capturing the impact of pricing on return behavior.

$$Uplift_{RF} = \left(\frac{\frac{RR_{after}}{GR_{after}}}{\frac{RR_{before}}{GR_{before}}}\right) - 1 = \frac{\frac{(\mathrm{UR_{after}} * \mathrm{AR_{after}})}{(\mathrm{US_{after}} * \mathrm{AS_{after}})}}{\frac{\mathrm{UR_{before}} * \mathrm{AR_{before}}}{(\mathrm{US_{before}} * \mathrm{AS_{before}})}} - 1$$

$$= \frac{RRate_{after}}{RRate_{before}} * \frac{\left(\frac{AR_{after}}{AS_{after}}\right)}{\left(\frac{AR_{before}}{AS_{before}}\right)} - 1$$

$$= [(Uplift_{RRate} + 1) * \frac{(Uplift_{AR} + 1)}{(Uplift_{AS} + 1)}] - 1$$

$$= Uplift_{RRate} * \frac{(Uplift_{AR} + 1)}{(Uplift_{AS} + 1)} + \frac{(Uplift_{AR} + 1)}{(Uplift_{AS} + 1)} - 1$$

$$= Uplift_{RRate} * C_{RRate \to RF} + C_{RRate \to RF} - 1$$

$$= C_{RRate \to RF} * (1 + Uplift_{RRate}) - 1 \tag{2.6}$$

where

$$C_{RRate \to RF} = \frac{Uplift_{AR} + 1}{Uplift_{AS} + 1}$$

## B.5 Return fraction before store openings

**Table B3** Return fraction (%) before store openings

| Stores | All zip codes | Control zip codes | Treatment zip codes | Treatment vs. Control | p value t test |
|---|---|---|---|---|---|
| Large Store 1 | 2.76 | 2.76 | 2.83 | 0.06 | 0.77 |
| | (4.37) | (4.22) | (4.62) | (0.22) | |
| Large Store 2 | 2.86 | 2.90 | 2.82 | -0.08 | 0.61 |
| | (4.33) | (4.38) | (4.28) | (0.15) | |

Notes.
Standard deviations are reported in parentheses.
*p<0.05, **p<0.01, ***p<0.001

**Table B4** Return fraction (%) before store openings after removal of outliers

| Stores | All zip codes | Control zip codes | Treatment zip codes | Treatment vs. Control | p value t test |
|---|---|---|---|---|---|
| Large Store 1 | 2.07 | 2.07 | 2.06 | -0.01 | 0.94 |
| | (2.62) | (2.60) | (2.66) | (0.18) | |
| Large Store 2 | 2.16 | 2.14 | 2.19 | 0.05 | 0.74 |
| | (2.67) | (2.62) | (2.73) | (0.14) | |

Notes.
Standard deviations are reported in parentheses.
*p<0.05, **p<0.01, ***p<0.001

### B.6 Calculating return fraction and return rate uplifts

Calculating return fraction and return rate uplifts involves interpreting the logit model coefficients in terms of the odds ratio and then converting this back to the return fraction (or return rate). The logit model can be expressed as: $logit(P) = intercept + \beta \cdot dummy$, where $logit(P) = \log(\frac{P}{1-P})$. To compute odds ratios $(\frac{P0}{1-P0})$, we can proceed as follows:

- When dummy=0, $logit(P0) = intercept \implies \frac{P0}{1-P0} = exp(intercept)$
- When dummy=1, $logit(P1) = intercept + \beta \implies \frac{P0}{1-P1} = exp\ (intercept + \beta)$

Next, we convert these odds ratios to probabilities:

$P0 = \frac{exp(intercept)}{1+exp(intercept)}$ and $P1 = \frac{exp(intercept+\beta)}{1+exp(intercept+\beta)}$

The uplift in terms of return fraction can be calculated as: $Uplift = \frac{P1-P0}{P0}$

This formula allows us to quantify the relative change in the return fraction (or return rate) due to the influence of the dummy variable.

Table B5 below includes the calculations for the intercept, $\beta$ coefficients, $P0$, $P1$, and the the return fraction ($Uplift\_RF$) and return rate ($Uplift_{RRate}$) uplifts for Large Store 1 and Large Store 2.

**Table B5** Calculating return fraction and return rate uplifts

| | **Return Fraction** | | **Return Rate** | |
|---|---|---|---|---|
| | Large Store 1 | Large Store 2 | Large Store 1 | Large Store 2 |
| Intercept | (3.447) | (3.447) | (3.399) | (3.399) |
| Beta | 0.158 | 0.142 | 0.249 | 0.196 |
| P0 | 0.031 | 0.031 | 0.032 | 0.032 |
| P1 | 0.036 | 0.036 | 0.041 | 0.039 |
| Uplift | 16.50% | 14.71% | 27.11% | 20.81% |

# Appendix C

## C.1 Experiment design and implementation

**Table C1** The list of physical stores in the study

| **Region** | **Stores** | **Opening year** | **Size** |
|---|---|---|---|
| Region A- North | Test Store 1 | 2008 | Large |
| | Store 1 | 2021 | Medium |
| | Store 2 | 2019 | Small |
| Region B - West | Test Store 2 | 2017 | X-Large |
| | Store 3 | 2021 | X-Large |
| | Store 4 | 2013 | Medium |
| | Store 5 | 2020 | Medium |
| Region C - South | Test 3 | 2008 | Medium |
| | Store 6 | 2019 | Large |
| | Store 7 | 2019 | Large |
| | Store 8 | 2021 | Large |
| | Store 9 | 2022 | Large |

**Figure C1** Timeline of events related to Shift Leads

| | | | | | | | | | ***Intervention (May 15-Jun 25)*** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nov | Dec | Jan | Feb | Mar 6th | … | Apr | May 1 (Mon.) | May 8 | May 15 (Mon.) | May 22 | May 29 | Jun 5 | Jun 12 | Jun 19 |
| SL role introduced | | | | SL dashboard introduced | | | | | *Wk 1* | *Wk 2* | *Wk 3* | *Wk 4* | *Wk 5* | *Wk 6* |

***Shift Lead Job description.*** The job description obtained from the retailer partner was translated to English.

As an Advisory and Service employee, you are responsible for astonishing customers in the store and through our other customer contact channels, from the moment they come in and need advice on a purchase, pick up a product or have a service issue.
You do that by:

- You find out customer demand with the help of an open attitude. Based on this customer question, you help them choose the best product or which solution we can offer.
- You advise customers about additional products, services and accessories so that they have everything they need when using the product.
- You solve technical problems for the customer on the spot where possible. If this is not possible, ensure correct intake of the product.
- You advise customers with a return so well that they leave the store with a better alternative.
- You amaze customers by personally approaching, welcoming and helping them.
- You follow your own performance and scores in the results card to get a little better every day.
- You make sure the store looks perfect so that you can be proud of it.
- You get a little better every day by following training courses in our own Study Factory. This way you maintain your knowledge and skills and you remain boss in the profession.
- You teach new Advice and Service employees to practice the profession in the Coolblue way.
- In addition to the customers in the store, you also help customers through our other customer contact channels in the moments when you are not helping a physical customer in the store.
- You are able to help anywhere in the store even if this falls outside your standard tasks, such as counting in the warehouse or changing store walls.
- You have an eye for the entire customer journey and propose improvement proposals to your manager.

***Shift Lead Dashboard.*** Below are the metrics available on the shift lead dashboard:

Worked hours and sales per worked hour:

- Time spent on customers
- Number of employees serving customers
- Number of employees not serving customers
- Total planned hours (including breaks)
- Total sales
- Worked hours
- Sales per worked hour
- Estimated sales vs. number of working employees
- Service level per hour

Visitors:

- Current average waiting time
- Current average transaction time
- Number of customers in queue
- Number of customers in service
- Service level percentage (all day)

Subcategories for visitor metrics:

- For Advice:
    - Current average waiting time for advice
    - Number of customers in queue for advice
    - Number of employees serving advice customers
- For Pickup:
    - Current average waiting time for pickup
    - Number of customers in queue for pickup
    - Number of employees serving pickup customers
- For Service:
    - Current average waiting time for service
    - Number of customers in queue for service
    - Number of employees serving service customers

***Real-time staffing adjustments tracking sheet template.*** The tracking sheet template was translated to English.

**Figure C2** The list of physical stores in the study

| Day | Date | Decision maker name | Up/down scaling | Time of the decision | Why? | What considerations did you make? | Which data (metrics) did you use? | What are you still missing/do you need to make better choices? (data, qualitative information etc) | Evaluation |
|---|---|---|---|---|---|---|---|---|---|
| Monday | 01-05-2023 | | | | | | | | |
| Tuesday | 02-05-2023 | | | | | | | | |
| Wednesday | 03-05-2023 | | | | | | | | |
| Thursday | 04-05-2023 | | | | | | | | |
| Friday | 05-05-2023 | | | | | | | | |
| Saturday | 06-05-2023 | | | | | | | | |
| Sunday | 07-05-2023 | | | | | | | | |
| *Monday* | *08-05-2023* | | | | | | | | |
| *Tuesday* | *09-05-2023* | | | | | | | | |
| *Wednesday* | *10-05-2023* | | | | | | | | |
| *Thursday* | *11-05-2023* | *You are part of a pilot and therefore up- and downscaling is not allowed on these days* | | | | | | | |
| *Friday* | *12-05-2023* | | | | | | | | |
| *Saturday* | *13-05-2023* | | | | | | | | |
| *Sunday* | *14-05-2023* | | | | | | | | |

***Roll-out Announcement.*** The announcement email was translated to English.

Hi!

From May 1, we have asked you, as part of the day management and margin per hour worked, to scale up and down and to log this. Nice to see that (many of) you do this too!

**Collaboration with University**

And that's good news! They will help us to link scaling up and down to margin per hour worked. That insight helps you to make an even better choice to scale up or down.

**What does this mean for scaling up and down?**

Not so much! The University does need data that they can compare with each other. That is why we are introducing test and control stores.

**Test stores**

- The test stores are
- From Monday 15 May you will have your own sheets for logging up and down. You will find the sheet of your store in this folder . You and your Shiftleads have access to the sheet
- As a test store, you are **not** allowed to scale up and down in some weeks . This is clearly stated in your sheet

**Control shops**

- These are the other shops;
- From Monday 15 May you will have your own sheets for logging up and down. You will find the sheet of your store in this folder . You and your Shiftleads have access to the sheet.
- Other than that, nothing changes in the process!

**1 more time. What do we expect from you?**

- That you actively manage scaling up and down as part of the day-to-day management. As a result, you have the right staffing and you contribute to the margin per hour worked
- You record in the sheet whether you have scaled up or down and indicate what you have taken into account
- Also look back a day and evaluate how you did. State what you have learned
- Stay on top of this as (Assistant) Store Manager and challenge your Shiftlead to contribute to this

**To ask?**

Next Monday, your Regional Manager will discuss the above with you during your weekly sales statement meeting. If you have any questions, please contact your regional manager.

## C.2 Intent-to-treat effect robustness checks

**Table C2** Week and day of the week fixed effects results

| Variable | log(Productivity) (1) | log (Labor hours) (2) | Log (Sales) (3) | log (Conversion) (4) | log (Basket value) (5) |
|---|---|---|---|---|---|
| $TREAT$ | 0.059** | -0.049 | 0.010 | 0.008 | 0.002 |
| | (0.028) | (0.033) | (0.016) | (0.011) | (0.015) |
| Week fixed effects | Yes | Yes | Yes | Yes | Yes |
| Day of the week fixed effects | Yes | Yes | Yes | Yes | Yes |
| $N$ | 2911 | 2911 | 2911 | 2911 | 2911 |
| $R^2$ | 0.679 | 0.863 | 0.920 | 0.370 | 0.525 |

*Notes.*
Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

**Table C3** Date x day of the week fixed effects versus day of the week fixed effects

| Variable | log (Productivity) (1) | log (Labor hours) (2) | log (Productivity) (3) | log (Labor hours) (4) |
|---|---|---|---|---|
| $TREAT$ | 0.060** | -0.049 | 0.058* | -0.048 |
| | (0.034) | (0.037) | (0.032) | (0.036) |
| Date x day of the week fixed effects | Yes | Yes | | |
| Week fixed effects | | | Yes | Yes |
| Day of the week fixed effects | | | Yes | Yes |
| $N$ | 2911 | 2911 | 2911 | 2911 |
| $R^2$ | 0.712 | 0.824 | 0.658 | 0.797 |

*Notes.*
Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

**Table C4** Exclusively experiment period

| Variable | log(Productivity) (1) | log(Labor hours) (2) | log(Sales) (3) |
|---|---|---|---|
| $TREAT$ | 0.036* | -0.036*** | -0.000 |
| | (0.019) | (0.013) | (0.016) |
| Date x day of the week fixed effects | Yes | Yes | Yes |
| $N$ | 502 | 502 | 502 |
| $R^2$ | 0.746 | 0.928 | 0.932 |

*Notes.*
Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

**Table C5** Excluding post experiment

| Variable | log(Productivity) (1) | log(Labor hours) (2) | log(Sales) (3) |
|---|---|---|---|
| $TREAT$ | 0.064** | -0.053 | 0.011 |
| | (0.031) | (0.036) | (0.017) |
| Date x day of the week fixed effects | Yes | Yes | Yes |
| $N$ | 2743 | 2743 | 2743 |
| $R^2$ | 0.736 | 0.881 | 0.934 |

*Notes.* Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

**Table C6** Excluding the first week

| Variable | log(Productivity) (1) | log(Labor hours) (2) | log(Sales) (3) | log (Conversion) (4) | log (Basket value) (5) |
|---|---|---|---|---|---|
| $TREAT$ | 0.059* | -0.048 | 0.011 | 0.009 | 0.002 |
| | (0.030) | (0.034) | (0.017) | (0.012) | (0.016) |
| $N$ | 2827 | 2827 | 2827 | 2827 | 2827 |
| $R^2$ | 0.731 | 0.884 | 0.933 | 0.551 | 0.588 |

*Notes.* Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

**Table C7** Excluding the local holidays

| Variable | log(Productivity) (1) | log(Labor hours) (2) | log(Sales) (3) | log(Conversion) (4) | log (Basket value) (5) |
|---|---|---|---|---|---|
| $TREAT$ | 0.058* | -0.047 | 0.012 | 0.008 | 0.004 |
| | (0.031) | (0.035) | (0.017) | (0.012) | (0.016) |
| $N$ | 2868 | 2868 | 2868 | 2868 | 2868 |
| $R^2$ | 0.732 | 0.881 | 0.933 | 0.557 | 0.593 |

*Notes.* Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

**Table C8** Times on-off-on

| Variable | log (Productivity) (1) | log (Labor hours) (2) | log (Sales) (3) |
|---|---|---|---|
| $TREAT$ x ($TIMES\_SCALED$=1) | 0.073** | -0.053 | 0.020 |
| | (0.029) | (0.041) | (0.029) |
| $TREAT$ x ($TIMES\_SCALED$=2) | 0.007 | 0.005 | 0.012 |
| | (0.057) | (0.062) | (0.038) |
| Date x day of the week fixed effects | Yes | Yes | Yes |
| $N$ | 2911 | 2911 | 2911 |
| $R^2$ | 0.732 | 0.882 | 0.933 |

*Notes.* Standard errors clustered at store × week level and reported in parentheses.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

## C.3 Potential instrumental variables and their definitions

**Ticket Forecast Deviation**

| | |
|---|---|
| Deviation from ticket forecast | $DevTICK_{it} = DevTICK_{it} - PlanTICK_{it}$ |
| Relative deviation from ticket forecast | $RelDevTICK_{it} = \frac{DevTICK_{it}}{PlanTICK_{it}}$ |

**Labor Forecast Deviation**

| | |
|---|---|
| Deviation from planned labor hours | $DevLBR_{it} = LBR_{it} - PlanLBR_{it}$ |
| Relative deviation from planned labor hours | $RelDevLBR_{it} = \frac{DevLBR_{it}}{PlanLBR_{it}}$ |

**Customer Waiting Time**

| | |
|---|---|
| Average customer waiting time (minutes) | $WAIT_{it} = \frac{1}{TICK_{it}} \sum_{j=1}^{TICK_{it}} wait_{ijt}$ where $wait_{ijt}$ is the customer wait time for ticket $j$ |
| Peak customer wait time | $MAX\_WAIT_{it}$ |
| Standard deviation of customer wait times | $SD\_WAIT_{it}$ |
| Coefficient of variation of wait times | $CV\_WAIT_{it} = \frac{SD\_WAIT_{it}}{WAIT_{it}}$ x 100 |
| Average of average customer waiting for the past 7-days | $MEAN_WAIT_{it}$_l1_to_l7 = $\frac{1}{7}\sum_{t=t-7}^{t-1} WAIT_{it}$ |
| Ratio of current to past average waiting times: | $RATIO_WAIT_{it}$_l1_to_l7 = $WAIT_{it}$_l1 / $MEAN_WAIT_{it}$_l1_to_l7 |

**Service Time**

| | |
|---|---|
| Average customer service time (minutes) | $SERVICE_{it} = \frac{1}{SERV_{it}} \sum_{k=1}^{SERV_{it}} service_{ijt}$ where $service_{ijt}$ is the customer service time for ticket $j$ |
| Standard deviation of customer service times | $SD\_SERVICE_{it}$ |
| Peak customer service time | $MAX\_SERVICE_{it}$ |
| Coefficient of variation of service times | $CV\_SERVICE_{it}$ = ($SD\_SERVICE_{it}$ / $SERVICE_{it}$) x 100 |
| Average of average customer service times for the past 7-days | $MEAN_SERVICE_{it}$_l1_to_l7 = $\frac{1}{7} \sum_{t=t-7}^{t-1} SERVICE_{it}$ |
| Ratio of service time to average for the past 7-days | $RATIO_SERVICE_{it}$_l1_to_l7 = $SERVICE_{it}$_l1 / $MEAN_SERVICE_{it}$_l1_to_l7 |

**Workload (ticket per labor hour)**

| | |
|---|---|
| Workload | $LOAD_{it} = \frac{TICK_{it}}{LBR_{it}}$ |
| Planned workload | $PlanLOAD_{it} = \frac{PlanTICK_{it}}{PlanLBR_{it}}$ |
| Deviation from planned workload | $DevLOAD_{it} = LOAD_{it} - PlanLOAD_{it}$ |
| Relative deviation from planned workload | $RelDevLOAD_{it} = \frac{DevLOAD_{it}}{PlanLOAD_{it}}$ |

**Workload (ticket per employee head count)**

| | |
|---|---|
| Workload | $LOADE_{it} = \frac{TICK_{it}}{EMP_{it}}$ |
| Planned workload | $PlanLOADE_{it} = \frac{PlanTICK_{it}}{PlanEMP_{it}}$ |
| Deviation from planned workload | $DevLOADE_{it} = LOADE_{it} - PlanLOADE_{it}$ |
| Relative deviation from planned workload | $RelDevLOADE_{it} = \frac{DevLOADE_{it}}{PlanLOADE_{it}}$ |

**Workload (tickets per a processing server)**

| | |
|---|---|
| Workload | $LOADS_{it} = \frac{TICK_{it}}{SERVER_{it}}$ |
| Planned workload | $PlanLOADS_{it} = \frac{PlanTICK_{it}}{PlanSERVER_{it}}$ |
| Deviation from planned workload | $DevLOADS_{it} = LOADS_{it} - PlanLOADS_{it}$ |
| Relative deviation from planned workload | $RelDevLOADS_{it} = \frac{DevLOADS_{it}}{PlanLOADS_{it}}$ |

**Occupancy**

| | |
|---|---|
| Employee occupancy rate per hour | $OR_{it} = (\sum_{k=1}^{SERV_{it}} service_{ikt}$ / $LBR\_floor_{it}) * 100$ where $service_{ikt}$ is the service time for ticket $k$ and $SERV_{it}$ denotes the total number of tickets served |
| Average occupancy over for the past 7-days | $MEAN_OR_{it}$_l1_to_l7 = $\frac{1}{7} \sum_{t=t-7}^{t-1} OR_{it}$ |
| Ratio of occupancy to average for the past 7-days | $RATIO_OR_{it}$_l1_to_l7 = $OR_{it_l1}$_l1 / $MEAN_OR_{it}$_l1_to_l7 |

**Employee head count**

| | |
|---|---|
| Store head count | $EMP_{it}$ |
| Floor staff head count | $EMP\_floor_{it}$ |
| Floor staff & management head count | $EMP\_floor\_mgmt_{it}$ |

**Other**

| | |
|---|---|
| Conversion | $CR_{it} = \frac{TRANS_{it}}{TICK_{it}}$ |
| Basket value | $BV_{it} = \frac{SALES_{it}}{TRANS_{it}}$ |
| Productivity | $PROD_{it} = \frac{SALES_{it}}{TICK_{it}}$ |
| Average productivity over for the past 7-days | $MEAN_PROD_{it}$_l1_to_l7 = $\frac{1}{7} \sum_{t=t-7}^{t-1} PROD_{it}$ |