THE IMPACT OF MOBILE APPS ON ONLINE AND OFFLINE SHOPPING BEHAVIORS

A Dissertation

by

UNNATI NARANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Venkatesh Shankar
Committee Members,	Fernando Luco
	Sridhar Narayanan
	Manjit Yadav
Head of Department,	Manjit Yadav

May 2020

Major Subject: Business Administration

Copyright 2020 Unnati Narang

ABSTRACT

In recent years, the penetration of mobile devices has reached unprecedented levels. Mobile apps account for 87% of mobile usage. Retail apps are among the fastest growing app categories. Do mobile apps influence shoppers' online and offline purchases and product returns? Further, do in-app experiences, such as app failures experienced by shoppers influence their purchases? The nascent literature on mobile apps is silent on these questions. In this research, I expand the scope of mobile marketing literature by quantifying and explaining the impact of mobile app introduction and app failures on shopping across channels. Unlike prior studies that focus on only online purchases, I consider both online and in-store purchases, and product returns.

In Essay 1, I quantify the impact of mobile app introduction on online and offline purchases and product returns. I leverage data on a large multichannel retailer's mobile app introduction and use a difference-in-differences regression. I find a 37% lift in net purchases in online and offline channels due to the launch of an app over 18 months. I outline important practical mobile marketing and cross-channel strategies for retailers.

In Essay 2, I quantify the impact of mobile app failures on online and offline purchases. I leverage exogenous systemwide failure shocks in a large multichannel retailer's mobile app and related data. Importantly, I investigate the heterogeneity among shoppers based on past relationship with the retailer. My results show a 7.1% decrease in purchases in stores due to an app failure. I outline important failure preventive and recovery strategies for app providers.

DEDICATION

This dissertation is dedicated to my parents.

ACKNOWLEDGEMENTS

This dissertation would not be possible without the unending support from my advisor and committee chair, Dr. Venky Shankar. I am thankful for his unconditional support and expert mentorship throughout the course of this research.

I am also most grateful to my committee members Dr. Manjit Yadav, Dr. Fernando Luco, and Dr. Sridhar Narayanan. I am thankful for the excellent training I have received in Marketing and in Economics during many seminars, conferences, job talks and other interpersonal interactions. I thank Alina Sorescu, Len Berry, Aric Rindfleisch, Avi Goldfarb, Nitin Mehta, Kathy Li, Dan McCarthy, Steve Puller, Jennifer Doleac, Andrew Barr, Jon Meer, Jason Lindo, Elena Fumagalli, and Yuna Choe and the staff at Mays Business School, especially John Norton. Much gratitude to Dean Eli Jones and Duane Ireland. Thanks also to TAMU Student Counseling Services and Rec Sports.

Thanks to my mom and dad for being my support and sounding board even from thousands of miles away. Thanks to my brother and his family for being there for me. Thanks to Pratyush who supported and relocated with me so I could start pursuing this research five years ago. Thanks to my cousins Raj and Shanu Pannu for planting the early seeds of a PhD in me, and my mentor Hitendra Wadhwa for inviting me to CBS. I thank my friends Anshuman, Swapnil, Sharon, Hari, Adam, Hannah, Cathy and others.

Most of all, I am grateful to my partner Josh, who has made this journey worth it and with whom, I am excited to share the fruits of the last few years of hard work most.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor(s) Venky Shankar (advisor) and Manjit Yadav of the Department of Marketing, Professor Fernando Luco of the Department of Economics, Texas A&M University, and Professor Sridhar Narayanan of the Department of Marketing, Stanford University.

The data analyzed was obtained through the Center for Retailing Studies, Mays Business School, Texas A&M University.

All data-related work conducted for the dissertation was completed by the student independently.

Funding Sources

There were no external sources of funding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

The retail industry in the United States (U.S.) contributes 5.9% to the U.S. gross domestic product and generates $1.4 trillion value. However, the retail environment is constantly evolving.

Historically, this unique and customer-facing industry has experienced several major changes. From the departmental stores of the mid 1800s-early 1900s, the malls and “big box” retailers of the 1970s, and the present-day online retailing, retail formats have adapted to changing customer lifestyles, needs, popular culture, and infrastructure. Historians that have documented the early rise of retail traditionally characterize retailers based on their size, scale, product categories, and locations. Indeed, the term “big-box” retailer was derived from the store’s physical appearance. Located in large-scale buildings of more than 50,000 square feet, such stores are designed in the shape of a box, e.g., Walmart, Best Buy and Ikea.

In the early days of retailing, innovations that offered retailers a competitive advantage were focused on the physical aspects of the stores. Retailers innovated by having creative displays, unique shelf space allocation, or discount events and clearance sales. In fact, some of the earliest innovations in retailing were targeted at making shopping an experience rather than a chore. In the words of British retailing maverick Harry Gordon Selfridge: *A shop should be like a song you never tire of.* The focal question for marketers was then: How can retailers bring customers to the store and make them spend more time there, and in turn, spend more dollars and share of wallet? The elements they could control were physical and the competition was largely local.

With the rise of the Internet in the 1990s though, retail went through tectonic shifts. Suddenly, the competition was global, and not local. Entry barriers were lower. Consumer choice exploded. Search costs reduced significantly. Overall, the Internet empowered consumers, who were better informed and could make find the "best" deal out there and still have it delivered to their doorstep when they wanted. No wonder, just over the last few decades, companies like Amazon have leveraged this massive opportunity. Ecommerce contributes over 40% to overall retail revenues now, and majority of those come from players like Amazon.

In parallel, traditional brick-and-mortar stores have had to shift their marketing strategy. First, there is a rise in brick-and-mortar stores opening online channels. Second, there is significant downsizing and resource allocation in the offline world, with retailers shutting down their loss-making or least profitable units. Finally, there is a greater thrust and need to understand cross-channel synergies online and offline. In some of the early research in cross-channel retailing in the Internet era academics highlighted online-offline synergies for retailers (Avery et al. 2012, Brynjolfsson et al. 2013).

The next big round of innovations in the post-internet era is the advent of smartphones (Shankar and Balasubramanian 2009). More than 2 billion people around the world own a smartphone. Much of this growth has come in the last decade as mobile technologies have sought to be one of the fastest-growing technologies of our times. Within the U.S., more than 80% adults use a smartphone. They spend a whopping 3 hours on average using these devices. Individuals' interaction and engagement with mobile devices opens up a huge opportunity for retailers to reach consumers anytime-anywhere. However, the impact of mobile technologies is not fully known or understood for large multichannel retailers that have both online and offline presence.

The focus of this dissertation is to examine the new technological innovation, mobile technologies, with respect to their impact on online and offline shopping behaviors. Within mobile technologies, mobile apps constitute the majority of consumer time spent. Mobile apps account for 87% of mobile usage. Retail apps are among the fastest growing app categories. Do mobile apps influence shoppers' online and offline purchases and product returns? Further, do in-app experiences, such as app failures experienced by shoppers influence their purchases? The nascent literature on mobile apps is silent on these questions. In this dissertation, I expand the scope of mobile marketing literature by quantifying and explaining the impact of mobile app introduction and app failures on shopping across channels. Unlike prior studies that focus on only online purchases, I consider both online and in-store purchases, and product returns.

In Essay 1, I quantify the impact of mobile app introduction on online and offline purchases and product returns. I leverage data on a large multichannel retailer's mobile app introduction and use a difference-in-differences regression. I find a 37% lift in net purchases in online and offline channels due to the launch of an app over 18 months. I outline important practical mobile marketing and cross-channel strategies for retailers.

In Essay 2, I quantify the impact of mobile app failures on online and offline purchases. I leverage exogenous systemwide failure shocks in a large multichannel retailer's mobile app and related data. Importantly, I investigate the heterogeneity among shoppers based on past relationship with the retailer. My results show a 5.5% decrease in purchases in stores due to an app failure. I outline important failure preventive and recovery strategies for app provider.

In the two essays in this dissertation, I examine the impact of mobile apps on online and offline shopping. Together, these essays quantify the impact of apps for large-scale retailers and provide prescriptive implications for such retailers to compete with giants like Amazon by leveraging emerging mobile technologies. Essay 1 highlights the *bright side* of retailers' branded mobile app introduction while Essay 2 uncovers a *dark side* of such app investments via app failures. Together, the two essays suggest that retailers expanding into new mobile technologies must maximize the gains from launching a new branded app while minimizing the potential risks due to failure.

The rest of this dissertation is structured as follows. In Chapter 2, I present an overarching view of the current literature in mobile marketing, particularly mobile apps with a view to set the stage for my essays and ways in which they contribute to the extant research. In Chapter 3, I focus on the impact of the launch of a branded retailers' mobile apps (Essay 1). In Chapter 4, I examine the impact of app failure in mobile apps (Essay 2). In Chapter 5, I make an effort to integrate the two essays. I conclude with key takeaways in Chapter 6.

CHAPTER II

AN OVERVIEW OF EMERGING LITERATURE IN MOBILE MARKETING*

The penetration of mobile devices has reached unprecedented levels. In 2017, 68.9% of the United States (U.S.) population or 224.3 million people owned a smartphone (Statista 2018). An average U.S. adult spends about five hours every day on his/her smartphone (TechCrunch 2017). Mobile apps are increasingly dominating mobile device use. Mobile apps account for 87% of mobile usage for an average adult in the U.S., which constitutes the bulk of digital media time (eMarketer 2017).

Given the rising popularity of mobile devices and apps, it is unsurprising that mobile marketing has become a strategic priority for firms (Shankar and Balasubramanian 2009; Shankar 2012). Mobile marketing refers to the two- or multi-way communication and promotion of an offer between a firm and its customers using a mobile medium, device, platform, or technology (Shankar and Balasubramanian 2009). For example, 51% of all digital ad budgets in 2016 were spent on mobile alone (IAB 2016). By 2018, mobile advertising spending in the U.S. will surpass TV advertising spending and account for 69.9% of all digital advertising spending, close to $40 billion (eMarketer 2018). Thus, mobile marketing is a topic of key interest for both academics and practitioners.

Research on mobile marketing has exploded in the last decade but has largely remained scattered. Prior reviews of mobile marketing research (Shankar and

---

*Reprinted by permission, Unnati Narang and Venky Shankar, Mobile marketing 2.0: State of the art and research agenda, *Marketing in a Digital World (Review of Marketing Research)*, 16, 97-119, 2019. Copyright (2019), Emerald Publishing Limited Howard House, Wagon Lane, Bingley BD16 1WA, United Kingdom.

Balasubramanian 2009; Shankar 2012) have provided a solid foundation to understand mobile marketing. With advances in mobile apps, augmented reality and wearables, mobile marketing has entered its second phase or Mobile Marketing 2.0. This phase exhibits three unique aspects. First, mobile device usage for digital media consumption has surpassed desktop usage in this new era. Second, the scope of mobile devices has expanded beyond smartphones and tablets to wearables and other smart devices such as smart speakers. Third, the integration and inter-connectedness of devices is more ubiquitous with the spread of Internet of Things (IoT).

Recent mobile marketing research has primarily addressed questions relating to the adoption of smartphones, tablets and apps, the effectiveness of targeting in mobile promotions, the influence of mobile display and in-app advertising, and the cross-channel impact of mobile devices and apps on offline behaviors in contexts such as retailing, healthcare, and gaming. The methods and data used to address relevant mobile marketing questions have become increasingly sophisticated and have leveraged large-scale field experiments, natural experiments, and structural models. Because mobile marketing is still nascent and evolving, and huge sums of money are being invested in mobile, there is potential to continue to gather new and unique data to better understand and predict the future of mobile-led marketing.

The Mobile Marketing 2.0 era also exhibits characteristics that are either new or not previously apparent. Emerging trends in app monetization, mobile augmented reality, data privacy and user experience, wearables, IoT, and self-driving cars, pose new questions for both practitioners and researchers.

In this chapter, we provide an overview of mobile marketing research, in particular, in the post mobile app phase. Based on the above discussion, we first outline insights from extant research along the three themes of targeted mobile promotions, personalized mobile advertising, and mobile-led cross-channel effects. These are areas where mobile marketing research has made significant advances in recent years. Next, we focus on underexplored emerging mobile marketing trends. We first identify and describe these new phenomena. Subsequently, for each area of extant and emerging mobile marketing theme, we outline a future research agenda for mobile marketing researchers for creating further knowledge in the Mobile Marketing 2.0 era.

Review of Mobile Marketing Research

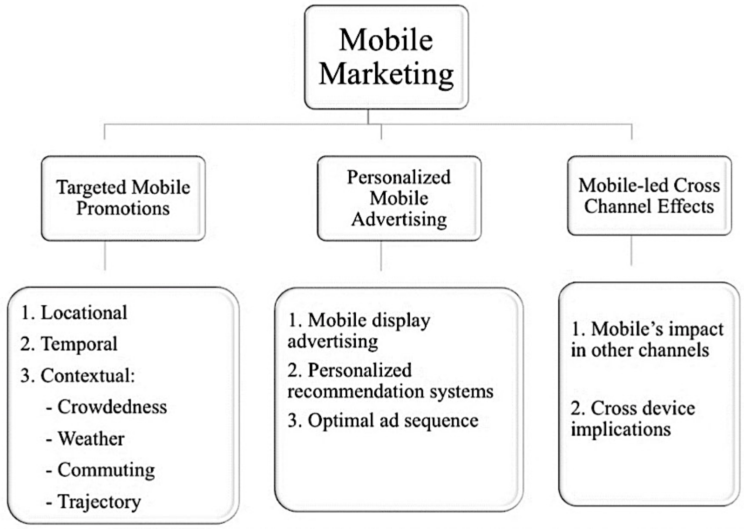
What makes mobile devices unique for marketers? It is useful to briefly review the unique characteristics and capabilities of mobile devices that lend themselves to marketing activities. Mobile devices are uniquely characterized by personability, portability, location-specificity, interactivity via touch interfaces, and compatibility with customer’s lifestyles and their use of other devices (Shankar and Balasubramanian 2009). Mobile devices have expanded to wearable devices, such as fitness trackers and virtual reality (VR) headsets. Mobile devices enable marketers to target customers based on their temporal, geographical, behavioral or contextual factors, personalize the message delivered to them, and engage in two-way interactions and touchpoints across different channels. These unique capabilities can be leveraged by marketers for improved mobile advertising, promotions, search, and shopping.

Initial research in mobile marketing focused on topics such as wireless e-commerce (Shankar, O'Driscoll, and Reibstein 2003), mobile advertising (Shankar and Hollinger 2007), mobile couponing (Dickinger and Kleijnen 2008), and mobile retailing (Shankar et al. 2010). The first large scale set of commercial apps came to market in 2008 when iPhone created the app store. However, research on mobile marketing through mobile apps was slow to take off. As shopper marketing quickly became a central practice for marketers, in particular, consumer packaged goods companies, mobile apps started to play an important role in shopper marketing, leading the emergence of mobile shopper marketing (for a review of mobile shopper marketing, see Shankar et al. 2016).

In the rest of this section, we provide an overview of marketing research in its second phase or mobile marketing 2.0. We focus on research that leverages the emerging capabilities of mobile marketing and focuses on three central themes: targeted mobile promotions, personalized mobile advertising, and mobile-led cross-channel effects. Targeted mobile promotions relate to the coupons and offers sent to the customers on their devices based on different types of targeting. Personalized mobile advertising refers to the use of mobile banner and display ads, and the personalization of ads displayed to each individual. Finally, mobile-led cross channel effects relate to the impact of mobile app and device usage on shopping in online and offline channels.

Figure 1 shows the broad dimensions of these three developments in mobile marketing research, while Table 1 reports how these areas emerge from the unique capabilities and benefits offered by mobile devices and apps.

Figure 1. Recent Advances in Mobile Marketing



Note: Reprinted with permission from Narang and Shankar (2019b).

Table 1. Mobile Attributes and Related Marketing Advantages

Mobile device attributes	Related mobile marketing advantages	Description
Portability, location-specificity, personability	Targeted mobile promotions	Because each individual typically carries a personal mobile device(s) all the time, mobile offers a unique opportunity to marketers to target at different times and locations.
Personability, interactivity	Personalized mobile advertising	Because each individual typically owns personal mobile device(s) that allow two-way interactions, mobile marketers can personalize the mobile experience for each user based on their specific preferences, past click-thru behavior, etc.
Anytime-anywhere access, interactivity	Mobile-led cross-channel effects	Because mobile devices can provide anytime-anywhere information benefits to shoppers about products and nearby stores as well as immediate access to shopping and exposure to deals and offers on-the-go, they can influence shopping in other channels.

Note: Reprinted with permission from Narang and Shankar (2019b).

Targeted Mobile Promotions

Mobile devices are portable and personable, making them suited for targeted promotional messages (Danaher et al. 2015, Hui et al. 2013). Mobile technologies allow marketers to target customers with promotions based on their specific location, time, and context. Mobile marketers can leverage several factors that influence consumers at the right moment with the right message through context (e.g., crowdedness), location, time, saliency, and historical shopping patterns or trajectory (Ghose 2017).

Contextual targeting is about reaching the customer at the appropriate "micro-moment." It is the sum of all factors, circumstances, and associated behaviors that guide decision-making (Ghose 2017). Customers are more likely to be influenced by contextual mobile marketing efforts. Customers respond more positively to in-app restaurant recommendations that are designed to be "in-the-moment" rather than expert-driven (Adamopoulos and Tuzhilin 2015). Commuters in crowded subway trains are twice as likely to respond to a mobile offer by making a purchase than those in non-crowded trains mainly because crowdedness makes people turn inward and immerse more in their handheld devices (Andrews et al. 2015). Even different types of commuting contexts can influence consumers differently. Commuters who travel to or from work are about three times as likely to redeem their first mobile coupon compared to non-commuters (Ghose et al. 2018). Commuters also differ from non-commuters in responsiveness to coupons, expiration periods, and time taken for coupon redemptions. Weather also influences customers in their purchases and coupon redemption decisions.

Purchase responses to promotions tend to be greater (smaller) and faster (slower) in sunny (rainy) weather (Li et al. 2017).

Locational targeting leverages the customer's location when a promotion is delivered through mobile devices. Customers are more likely to make a purchase in response to a promotional offer received when they are closer to the store than when they are farther from it (Luo et al. 2013). Furthermore, location-based targeting is used as a tool to gain competitive advantage and/or to drive shoppers away from the competition to the focal firm. *Competitive* locational targeting is the practice of sending promotions to customers when they are near a competitor's location. It results in greater returns to promotional discount depth than targeting customers when they are close to the focal firm's location (Fong, Fang, and Luo 2015). Thus, mobile promotions have a competitive influence, in particular, when firms target based on historic, real-time location, and peak vs. off-peak time. Furthermore, the impact of mobile promotions based on such "geoconquesting" is low when the competitor also engages in similar targeting (Dubé et al. 2017). *Co-locational* targeting further incorporates customer's network structure based on the idea that an individual's location reveals the underlying preference for a type of place and event, and that co-located individuals who are at the same place at approximately the same time reflect similar preferences. Indeed, there is a significant relationship between co-location and response to mobile coupons in the same product category (Zubcsek, Katona, and Sarvary 2017).

Temporal targeting involves sending timely messages to customers based on real-time targeting. Promotions that temporally target customers increase their unplanned spending (Heilman, Nakamoto, and Rao 2002). The effects of temporal targeting are less straightforward to analyze when it is combined with locational targeting. For mobile users close to the focal firm's location, there is a negative sales-lead time relationship. Thus, sending them promotions on the same day is more effective than sending them promotions two days prior to the promoted event. For non-proximal mobile users, there is an interesting inverted-U relationship between response and temporal targeting. For example, large scale field experiments show that targeting non-proximal users with a one-day prior promotion is more effective than both same-day and two-days prior targeting (Luo et al. 2013). Thus, the interactions between the different types of targeting efforts result in different implications for customers and marketers.

Salience or position effects is an important consideration for mobile marketers in targeting and search. Mobile devices have a smaller screen size than desktops and laptops. As a result, ranking effects are higher on mobile phones suggesting higher search costs. Because of these search frictions, links that appear at the top of the screen are more likely to be clicked on mobile phones than on desktops and laptops (Ghose, Goldfarb, and Han 2012). In many settings, such as Yelp and other recommendation-based apps, user-generated content influences the formation of consideration sets through search. In such a setting, the sizes of mobile shoppers' consideration sets may be small due to search frictions. If a lower-ranked item in a search is not a part of the shopper's consideration set, it can result in sub-optional choices and decision-making (Muir and Tsai 2016). Overall, when people search on mobile, it tends to lead to action:

92% of those who searched on their phone made a related purchase, and 68% of people use search to enable decisions at some point in the future (Google 2018).

Finally, shoppers' historical movement patterns and offline trajectories provide rich insights into *how* they reach a specific location and what experiences might shape their subsequent shopping choices and coupon usage. A large-scale field experiment in one of Asia's largest malls leverages this idea of offline walking patterns or trajectories' impact on the redemption rates of mobile coupons (Ghose, Li, and Liu 2017). Their results show that such targeting leads to increased redemption probability, faster redemption behavior, and higher transaction amount. This effect is weaker on weekends when shoppers are more likely to be in exploratory shopping mode and less impulsive than during weekdays (Ghose et al. 2017).

Mobile influences shopping within the store (e.g., Grewal, Ahlbom, Beitelspacher, et al. 2018) show that mobile phone use can increases point-of-purchase sales through eye-tracking technology used in a field experiment. Mobile messages divert consumers from their conventional shopping loop and induce them to spend extra time in the store examining products and shelf prices.

Personalized Mobile Advertising

Mobile advertising includes banner/display advertising, search (paid and unpaid) advertising, and personalized advertising. Research on mobile advertising (e.g., Grewal et al. 2016) offers a framework in which contextual and consumer factors determine advertising goals, leading to the choice of advertising elements and advertising outcomes, such as clicks and purchases. Market and firm factors moderate the relationship between advertising goals and advertising elements.

Mobile devices are uniquely tied to an individual and allow marketers to leverage data on individual preferences, movement patterns, co-located social connections, and other individual-specific variables to personalize advertising and marketing communication messages. Thus, mobile devices offer a unique marketing opportunity for personalization. In the mobile advertising context, mobile display advertisements result in favorable attitudes towards products and purchase intentions, in particular, for high-involvement and utilitarian products (Bart, Stephen, and Sarvary 2014). The use of mobile ads with other communication channels (e.g., television) can result in a "tech mix" that facilitates information retrieval. In general, such ads are more effective when viewers are close to purchase and are highly involved and invested in the decision-making process.

Furthermore, with personalization, advertisements can be shown uniquely and/or in unique sequences to each user. There has been at least one recent effort in marketing research to develop a personalized dynamic framework for optimal ad allocation and sequence in which ads are shown to a mobile user (Rafieian and Yoganarasimhan 2018). The rationale for incorporating past ad exposures is to better understand user response to a current ad and user app usage. This is particularly true because seeing a sequence of ads within a short duration can influence users differently relative to seeing a sequence of ads over a long duration. Therefore, understanding the impact of previous sequence of ad exposures and ways of optimizing ad placement strategy within a session is critical in the mobile contexts. Furthermore, it requires personalization at the individual user level, since users are heterogeneous in exposures and their responses.

Personalization also captures customer heterogeneity, an increasingly important theme for enriching the understanding of different user experiences in mobile interfaces. Different users interact with and are influenced by mobile devices differently. A recent working paper shows that retail app users who are more loyal, have greater prior digital channel use, and who experienced failures less attributable to the retailer, are less sensitive to failures and crashes in the app than other retail app users (Narang, Shankar, and Narayanan 2018).

Furthermore, the interface and combination of interfaces on which the ads are displayed also affects shopper responses. Combining both web and mobile display advertising triggers more clicks and purchases than using only web or only mobile for display advertising (Ghose, Han, and Park 2013). Within mobile devices such as smartphones and tablets, ads can be delivered through apps using in-app advertisements, text messages or mobile browser webpages. The effects through these different interfaces might differ. The use of SMS (short messaging service) for advertising leads to higher awareness generation, brand attitudes, and direct behavioral responses (Barwise and Strong 2012).

Finally, mobile advertising strategies may operate differently for different cultures and countries. In emerging economies, a popular mobile marketing strategy is Missed Call Marketing (MCM), where the receiver can indicate a yes or no response with a free missed call (Knowledge @ Wharton 2016). It is also known as miskol in the Philippines, beep in Africa, memancing in Indonesia, and flashcall in Pakistan. A missed-call based campaign typically includes a call-to-action that requests a missed call back if the recipient wishes to get more information. The low cost and fast speed of missed-call campaigns are central to their appeal, in particular, in emerging and low-income economies. However, the impact of MCM is a relatively less researched area.

Mobile-led Cross-channel Effects

Mobile devices not only impact shopper behavior within the channel, but also have wider cross-channel implications (Shankar et al. 2016). In omnichannel retailing, mobile devices can facilitate purchases in stores and online. The unique characteristics of mobile devices that make this possible are anytime-anywhere accessibility, location-specificity and interactivity. Shoppers can glean product-, offers- and store- related information from apps (e.g., nearest store locator, stores open hour, latest products on offer). These capabilities enhance convenience for shoppers, the ease of finding relevant information (Dubé et al. 2017, Fong et al. 2015), and the ability to quickly shop and checkout (Shankar 2012, Shankar and Balasubramanian 2009). Many apps are linked with the retailer's loyalty programs and can provide shoppers easy access to their loyalty benefits across channels (e.g., Macy's and Kohls apps) as well as allow searches for rewards redeemable across channels.

Both mobile devices and apps have unique cross-channel effects relating to shopper purchases. Cross-device effects of mobile devices for m-commerce include effects on frequency, quantity, and monetary value of purchases in other devices. In a study of the introduction of Taobao's tablet app, Xu et al. (2016) find that overall ecommerce revenues increase after the tablet app launch. They also show that tablet commerce substitutes desktop commerce but complements smartphone commerce. Thus, the effects of mobile device addition to the shopper's purchase devices are not straightforward as they can lead to both synergies and cannibalization. Shoppers who transact more using a mobile device than a desktop computer purchase more frequently but spend less in monetary value (Lee et al. 2016).

Within mobile devices, apps offer a unique opportunity to create and enhance shopper loyalty. In the ecommerce context, eBay's app is associated with higher aggregate revenues from the platform (Einav et al. 2014). This is true even for apps that provide access to multiple retailers, such as air miles loyalty rewards apps, that let people redeem their points through various retailers. In such apps, shoppers' point accruals increase once they start using the app and two app features, information lookup and check-ins contribute most in increased shopper activity (Kim, Wang, and Malthouse 2015). Even promotional campaigns to encourage shoppers to use mobile apps result in higher spending once shoppers adopt ecommerce (Wang, Malthouse, and Krishnamurthi 2015). For non-retail settings, such as news apps, shoppers' website visits increase after the mobile channel's introduction (Xu et al. 2014).

While most studies examine the impact of mobile channel addition or adoption on online or ecommerce revenues, the impact on offline or in-store outcomes is less well-known. Mobile apps drive purchases in both online and in store channels (Narang and Shankar 2016). In-store effects are largely explained by the use of the app in close geographical-temporal proximity, in particular, for features, such as reward coupons discoverable through the app that can be redeemed in store and product discovery through the app. The app drives purchases of a larger diversity of products, in particular, less popular products. Despite higher purchases, app adopters are also likely to return more products than non-adopters after the app is introduced.

Marketers can utilize mobile devices in conjunction with retail kiosks to influence shopping. Such combination can be implemented through retail mobile integrated kiosks to allow marketer's information to flow seamlessly across the kiosk, the retailer's website, and mobile apps. Using both field experiments and lab studies, Grewal, Ahlbom, Noble, et al. (2018) show that inspirational communication content (e.g., food recipe) increases unplanned spending and sales more than promotional communication content (e.g., coupons). They further reveal that these effects are mediated by the activation of category-related thoughts and purchases of substitute products related to inspirational communication content and that the effects are greater for shoppers with low budget and for those who process information concretely. Their findings suggest that retailers should enhance sales by offering inspirational ideas to shoppers through mobile integrated kiosks mainly for low-budget, frequent, concrete processing shoppers.

A summary of the three emerging mobile marketing themes, their representative studies, and key findings appears in Table 2. These research studies also point to the emergence of new trends in mobile marketing.

Table 2. Summary of Key Issues in Mobile Marketing

Key Issue	Representative Studies	What We Know – Current Evidence
Targeted mobile promotions	Andrews et al. (2016) Danaher et al. (2015) Dubé et al. (2017) Fong et al. (2015) Ghose et al. (2018) Hui et al. (2013) Luo et al. (2013) Li et al. (2017) Zubcsek et al. (2017)	Mobile promotions targeted based on user's context (e.g., crowdedness, weather), location, time, saliency, and historical shopping patterns or trajectory impact coupon redemption rates, time, etc. depending on the type of targeting and their different combinations. Mobile promotions have a competitive influence, in particular, when firms target based on historic, real-time location, and peak vs. off-peak time
Personalized mobile advertising	Bart et al. (2014) Grewal et al. (2016) Rafieian and Yoganarasimhan (2018)	Mobile display advertisements result in favorable attitudes towards products and purchase intentions, in particular, for high-involvement and utilitarian products. Personalization and optimal mobile ad sequencing can further improve user response to ads. Customer heterogeneity is relevant for mobile marketing. Retail app users who are more loyal, have greater prior digital channel use, and who experienced failures less attributable to the retailer, are less sensitive to negative in-app experiences, such as app crashes and failures.
Mobile-led cross-channel effects	Einav et al. (2014) Grewal, Ahlbom, Beitelspacher et al. (2018) Grewal, Ahlbom, Noble et al. (2018) Kim et al. (2015) Lee et al. (2016) Narang and Shankar (2016) Narang et al. (2018) Wang et al. (2015) Xu et al. (2016)	Mobile apps and devices impact outcomes in online and offline channels. Specifically, mobile app adoption results in higher frequency, quantity and monetary value of purchases in online and offline channels, as well as higher product returns. Tablet commerce substitutes desktop commerce but complements smartphone commerce.

Note: Reprinted with permission from Narang and Shankar (2019b).

Emerging Trends

In what new ways does mobile marketing influence customer preference and decision-making? There is growing academic research on this issue, but it is still nascent to fully capture all the emerging trends and issues. In this section, we identify and describe seven emerging trends, consistent with our characterization of the Mobile Marketing 2.0 era. These trends relate to developments in existing mobile technologies (e.g., app monetization, augmented reality in apps), introduction of new smarter mobile technologies (e.g., self-driving vehicles, wearables, Internet of Things), and the engines powering these developments (e.g., big data, artificial intelligence). These trends are consistent with industry reports that offer an early glimpse into customer perceptions and behavior in an increasingly mobile-first world. We next delve into each of these emerging trends.

Mobile App Monetization

Mobile apps offer engagement, convenience, easy access, and shopping all through a single tap. In the past two years, mobile app downloads have surged by 60 percent globally (App Annie 2018). Mobile marketing is the "primary tool for the digital omnivore" and we are now in the "app age" with mobile apps becoming the hottest emerging trend for engaging users beyond one-off promotional redemption (comScore 2017).

An important emerging trend in mobile apps is app monetization. Global consumer spending for apps has now reached $86 billion. Several app categories are increasingly popular sources of revenue for app providers. These categories include ride-sharing, bike sharing, home rentals, health tracking, pro-social activity, and creative apps. In-app purchases (IAP), in-app advertisements (IAA) and paid apps are three formats for monetizing apps (Kumar, Dogan, and Lahiri 2016). App retailers such as Apple and Google incentivize apps that monetize content through subscription sign-ups and renewals. Finally, apps also provide benefits that influence spending and expenses outside of the app. For example, the adoption of an mHealth platform significantly reduces blood glucose and glycated hemoglobin levels, hospital visits, and medical expenses (Ghose, Lou, and Li 2018). Similarly, push notifications in apps can influence donation decisions and donation amounts (Lee, Gopal, and Lee 2017).

Mobile Augmented Reality (AR)

Augmented reality (AR) is the confluence of virtual reality with reality. Mobile technologies, such as built-in camera, sensors, and computational resources have made AR apps possible on mobile devices. Apps that leverage these technologies can generate new revenues. Estimates suggest that AR apps could boost iPhone and App Store sales by as much as $8 billion (CNBC 2018a). A popular AR app is Nintendo's Pokemon Go game that lets users chase game characters in public areas based on projections from their phone screens. While mobile AR apps engage users and integrate the online and offline environments, they can also have adverse consequences, including road accidents (Faccio and McConnell 2017).

In addition to gaming, mobile AR is quickly catching on in retail. For example, Amazon debuted AR View, its augmented reality shopping tool. AR View allows shoppers to see before buying how items will look in their homes (Engadget 2018).

Apart from improving shopping experience and satisfaction, AR tools can also potentially improve brand engagement and loyalty. For example, the wine brand 19 Crimes offers an AR app for free. Once users download the free app (iOS and Android), they can hold it up to the label on the wine, which features an image of a former convict. The convict comes alive with animation and narrate their story (Forbes 2017). IKEA offers a mobile app that lets users overlay the furniture in the store against their home picture to visualize how the furniture will appear in their own homes.

Data Privacy and Mobile User Experience

While mobile devices and apps improve user experience by making it more relevant and engaging, the algorithms that make this possible rely on huge amounts of user data. Apps elicit different types of data from their users (Kesler, Kummer, and Schulte 2018). Data breaches and mishandling of customer data have led to more stringent regulatory oversight. European Union's introduction of general data protection regulations (GDPR) in May 2018 puts ownership of customer data in the hands of users and requires firms to seek permission from users to use their data and be responsible custodians of customer data. These regulatory changes impact the way companies collect, store and use data, as well as how transparent they are in their communications with the users about these data-handling processes. The California consumer privacy act of 2018 is likely to reshape how companies handle data in the most populous state of the U.S. (Wakabayashi 2018). Under this law, firms will be required to disclose the data they collect from users and to let users decide whether their data can be sold.

These regulations raise a number of interesting research questions. What do these and similar regulations imply for mobile marketers? First, marketers need to invest in strengthening opt-in messaging, e.g., highlighting a benefit of the app before asking users to allow push notifications. The New York based health app, ZocDoc, does this by letting patients know that they can be updated of their appointments with doctors if they sign up for notifications before the users are shown the push notification prompt from the App Store/ Google Play. Personalized content in notifications can deliver a four-times lift in open rates vs. generic content and leveraging behavioral-based user actions can boost open rates by nine times (Leanplum 2017). However, opt-in can mitigate the privacy concerns surrounding notifications.

Wearables and Other Smart Devices

Another emerging trend and area of future investigation for mobile marketers is the ability to leverage the connections between mobile and other smart devices (Shankar et al. 2016). For example, using the Amazon Alexa app, consumers can utilize voice-based devices, such as Amazon echo, or echo dot to ask Alexa to perform different tasks, including shopping. Amazon has also tied up with hotels, such as Marriott by placing Amazon Echo speakers in hotel rooms so that guests can use voice for different services (Forbes 2018). Amazon creates engagement with Alexa by emailing users specific commands to use to get their voice assistant to perform certain tasks (e.g., “Alexa, reorder deodorant” or “Alexa, track my order”). Other ways apps link with wearables for improved user experience and convenience include recording and presenting archived data from tracking devices. For example, smart watches like Fitbit that track health data link back to the app to allow users to check their statistics at any time and enjoy additional features, such as connecting with social networks and friends who use the same device. In this way, these devices leverage both direct and indirect network effects.

Self-driving Vehicles

An important development in recent years is the emergence of self-driving vehicles, with several companies testing these vehicles on the roads. Self-driving cars are expected to hit the roads in the next five years. Estimates suggest that by 2035, 12 million fully autonomous cars would be sold (BCG 2017). This prospect would open up huge opportunities and challenges for marketers. Because marketers would be able to communicate with individual cars, the cars would serve as personalized advertising platforms, vehicles for delivery, and as a marketing channel (Gelb 2017).

Self-driving cars would provide marketers rich customer data and the ability to reach them directly and personally. Marketers can map and leverage customers' interactions with cars at a granular level. Furthermore, self-driving cars will offer customers greater convenience. Already, large-scale retailers, such as Kroger are teaming with self-driving car manufacturers and robotics firms to test these technologies for supplying groceries directly to their customers' homes. In addition, self-driving vehicles would have spillover effects on other industries, including ride-hailing and car sharing services (Belk 2014). In this way, autonomous or self-driving cars will likely present new challenges for existing auto-related services but also create new opportunities for industries, such as retailing.

Internet of Things (IoT)

The Internet of Things (IoT), a network of gadgets, devices, appliances, vehicles, and sensors embedded in objects, all connected to the Internet, is shaking up marketing. The network collects and exchanges data from and with all these items. As such, these items are mobile and serve to further expand the potential of mobile marketing. Marketers can use the network to better engage the customers in their journey. They can also harness customer interactions with all the items in the network in real time and make more effective decisions on the fly. The IoT can be viewed as an endless highway of touchpoints that can feed data on customers' digital footprints everywhere. Naturally, IoT provides a treasure trove of data on a myriad of customer activities.

The IoT can redefine marketing in different areas such as repeat purchases and customer services. For example, a washing machine connected to the IoT ecosystem can automatically reorder laundry detergent, and an inkjet printer on the IoT can replenish refills in advance, skipping customers' need recognition stage in their shopping journey. In the customer service realm, KONE, an elevator and escalator company, analyzes the data collected from the sensors in its elevators to anticipate symptoms and addresses them through technicians before they become problems (*Kone* 2018). The connectivity of multiple devices to the Internet through a giant network gives rise to the development of innovative services that could be aptly termed Internet of Services (IoS).

Artificial Intelligence (AI)

Artificial intelligence (AI), "programs, algorithms, systems or machines that demonstrate intelligence, or more generally, a set of tools that can enhance the intelligence of a product, service, or solution," is rapidly growing in importance and is reshaping retailing (Shankar 2018) and marketing, in general. AI is helping marketers leverage data to better understand and anticipate customer needs and make optimal decisions to maximize customer lifetime value. AI involves the use of machine learning models developed on big data (voluminous data from a variety of sources collected at high velocity) to predict customer behavior. Much of the data comes from mobile interactions of customers. In its most ubiquitous form, AI is embedded in the virtual assistants of smartphones (e.g., Siri, Cortana, Alexa) and smart speakers (e.g., Echo, Google Home, Apple Homepod), all mobile devices. AI is also assisting in specialized marketing decisions such as salesforce planning (e.g., Einstein) and retailing (e.g., IBM Watson).

Many firms are already using AI to improve and automate some of their marketing decisions. For example, French retailer, L'Occitane, uses AI to analyze customer data on its desktop and mobile websites and to personalize site layout, having experienced a 159% jump in mobile conversions on its U.K. site (Sandler 2018). Thank God It's Friday (TGIF), a U.S. restaurant chain, uses AI to integrate data from mobile app, email, loyalty program, and in-store purchases, and to make personalized offers to its customers through text messages (Marshall 2018). Thanks to their efforts over a year, TGIF's customer engagement multiplied by a factor of five and sales grew by $150 million (Marshall 2018).

A Research Agenda

With a better understanding of the ways in which academics have explored the evolving landscape of mobile marketing interventions, including those related to targeting, personalization, and cross-channel spillovers, we next develop an agenda for future research in mobile marketing. We incorporate central ideas from mobile marketing trends to recommend future research questions along the themes of current and emerging interest.

Research Agenda based on Gaps in Extant Literature

While mobile marketing research has advanced in the three areas of mobile targeting, personalization, and mobile-led cross-channel effects, there are several unexplored questions worthy of future investigation.

Research on targeted mobile promotions has examined the effects of locational, temporal, and contextual targeting on the immediate response to one-time promotions (Andrews et al. 2016; Fong et al. 2015). However, only a few studies have considered the impact of combining two or more different targeting strategies (Luo et al. 2013). Therefore, there is a need for further research on the effectiveness of different targeting approaches when used together. Opportunities also exist for understanding newer ways of targeting not explored before. Thus, two potential future research questions are: (1) What are other innovative and unexplored ways of targeting mobile promotions (e.g., based on dissimilarity between consumers, reinforcement learning based on past redemption behaviors)? and (2) How do various types of targeting work when implemented together?

Research on personalized mobile advertising is limited to one or two applications on developing personalized recommendation systems or personalizing the sequence of ads shown to users (Rafieian and Yoganarasimhan 2018). However, much work remains to be done in this area. Three potential ideas for future research are: (1) What are some effective ways of optimizing both ad sequence and ad content in mobile web and apps for maximizing user exposure, engagement, and actions? (2) What is the impact of personalization in different formats, such as in-app personalization, recommendations, and push notifications on purchases? and (3) How should marketers allocate mobile ad spending across different personalization strategies?

Research on mobile-led cross-channel effects, has assessed the impact of mobile devices and apps on primarily online outcomes (Wang et al. 2015, Xu et al. 2016). Not much is known about how mobile augments exposure to shopping in other channels or messaging from other channels. Furthermore, previous research has examined data from a single retailer or app provider. Thus, three potential areas of future research are: (1) How do different apps and/or portfolio of apps drive customers' offline choices? (2) What is the impact of mobile web vs. app on shopper preferences and behaviors? and (3) What are effective ways to optimize mobile shopping experiences and create synergies with other channels? Some illustrative potential research questions based on gaps in extant literature appear in Table 3.

Table 3. Research Agenda Based on Analysis of Gaps in Extant Literature

Key Issue	Future Research Agenda
Targeted mobile promotions	What are other innovative and unexplored ways of targeting mobile promotions (e.g., based on dissimilarity between consumers, reinforcement learning based on past redemption behaviors)? How do various types of targeting work when implemented together?
Personalized mobile advertising	What are some effective ways of optimizing both ad sequence and ad content in mobile web and apps for creating maximum user exposure, engagement, and actions? What is the impact of personalization in different formats, such as in-app personalization, recommendations, and push notifications on purchases? How should marketers allocate mobile ad spending across different personalization strategies?
Mobile-led cross-channel effects	How do different apps and/or portfolio of apps drive customers' offline choices? What is the impact of mobile web vs. app on shopper preferences and behaviors? What are some ways to optimize mobile shopping experiences and create synergies with other channels?

Note: Reprinted with permission from Narang and Shankar (2019b).

Research Agenda based on Emerging Trends

While industry is fast embracing the emerging trends in mobile marketing, particularly led by behemoths like Amazon, there is huge untapped opportunity for academics to understand newer forms of interaction and engagement that augment mobile marketing strategies.

Research on mobile app monetization can focus on several unanswered questions. Firms follow different app monetization approaches, ranging from free apps with in-app purchases to freemium apps with paid subscription for upgrades (Kumar et al. 2016). Which revenue models work best under what scenarios and at what prices? Some pending questions in the app monetization domain include: (1) What are some ways of effectively monetizing mobile apps? (2) How can providers create mobile app stickiness and engagement in the long term? and (3) What in-app features and functionalities do app users value based on heterogeneity in user and app characteristics?

Mobile AR (Augmented Reality) has allowed another new breakthrough in engagement by blurring the lines between on-screen and off-screen experiences (Shankar et al. 2016). Because AR requires significant time and effort, in this respect, firms and researchers alike would benefit from understanding: (1) What is the impact of AR apps on user engagement and actions in and outside of the app? (2) How do AR apps impact firm's brand image? What are some ways in which firms can leverage such apps in different settings and industries? and (3) What are the implications of combining AR with other non-mobile devices and settings (e.g., wearable in stores)?

In the mobile privacy and data sensitivity context, it is unclear how the new capabilities of mobile devices that leverage rich user-level data will attract and invite regulatory purview and concerns for providers. It is also unclear how firms should communicate the data handling processes in a simple but transparent manner with end users when eliciting information (Kesler et al. 2018). Several unanswered but pressing concerns can be delineated in this area: (1) What is the impact of stricter privacy regulations on mobile marketing? (2) Are there trade-offs associated with handling data privacy concerns and optimizing user experiences to be more relevant? What are some optimal approaches? (3) How do customers vary in their privacy preferences and how can marketers glean and leverage this information for creating unique mobile experiences for them? and (4) What is the broad spectrum of privacy for mobile users (e.g., different levels of data sharing)? Does the effect of different levels of privacy on user experience and sensitivity to promotions and advertising vary? How?

Another emerging trend in mobile marketing is cross-device integration with wearables and smart devices, such as smart speakers and watches (Shankar et al. 2016). In this domain, the key questions of interest are: (1) What are the future avenues for the growth of voice-enabled assistants? (2) How are voice-enabled assistants likely to impact different industries, e.g., retail and hospitality? (3) What are the ways in which mobile devices and apps can create and sustain synergies with other devices, including wearables? (4) Are there any potential cannibalization effects? and (5) What do we know about the usage patterns across devices?

From both marketing and policy perspectives, research on self-driving cars can address some unanswered questions (Waldrop 2015). Early advances in research relating to self-driving cars suggest that their broader adoption will depend on their price, ownership (individual or service companies), and liability for accidental damage. Potential research questions for marketing researchers are: (1) What are the effective ways of marketing self-driving cars? (2) What will be the impact of driverless cars on directly- and indirectly- related markets, such as ride sharing and leasing? and (3) In what ways can retailers leverage self-driving car technologies for improved delivery? Research on IoT is still in its infancy but can rapidly evolve with the availability of increased data from the ever-growing number of mobile devices connected to the IoT ecosystem. Analysis of such data can lead to more accurate predictions of customer behavior. In this regard, some important research questions are: (1) What models can help marketers deliver timely, personalized messages and offerings to customers based on customers' stages in their shopping journeys for different products?, (2) How can marketers overcome privacy issues involved in accessing multiple devices to collect relevant data and provide useful services to customers?, and (3) In what ways can marketers seamlessly develop Internet of Services (IoS) by leveraging the data from IoT?

Research on AI can have a fundamental impact on marketing. The proliferation of AI in consumers' daily lives and marketing decision makers' arsenal opens up an exciting new research frontier in marketing. Researchers are developing newer models for personalization that can form the engine behind AI recommendation systems (e.g., Darani and Shankar 2018, Jacobs, Donkers, and Fok 2016). Consistent with Shankar (2018), we identify three broad research questions of importance. (1) How do we develop AI-driven, mobile-enabled marketing systems for an array of marketing decisions ranging from product portfolio to pricing to promotion to personalization? (2) How do we develop new models of shopping behavior of AI- and mobile-assisted shoppers and consumers? and (3) How can we apply neuroscience-based AI models that leverage mobile data to develop a deep understanding of consumer psychology? The search for answers to these broad questions will spur new research on narrower and focused questions that can be addressed using the latest advances in data, machine learning, causal modeling, and experimental research.

Illustrative potential research questions based on emerging trends in mobile marketing appear in Table 4. These questions are grouped by the emerging trends.

Table 4. Research Agenda Based on Emerging Trends

Key Issue	Future Research Agenda
Mobile app monetization	What are some ways of effectively monetizing mobile apps? How can providers create mobile app stickiness and engagement in the long term? What in-app features and functionalities do app users value based on heterogeneity in user and app characteristics?
Mobile Augmented Reality (AR)	What is the impact of AR apps on user engagement and actions in and outside of the app? How do AR apps impact firm's brand image? What are some ways in which firms can leverage such apps in different settings and industries? What are the implications of combining AR with other non-mobile devices and settings (e.g., wearable in stores)?
Data privacy and mobile user experience	What is the impact of stricter privacy regulations on mobile marketing? Are there trade-offs associated with handling data privacy concerns and optimizing user experiences to be more relevant? What are some optimal approaches? How do customers vary in their privacy preferences and how can marketers glean and leverage this information for creating unique mobile experiences for them? What is the broad spectrum of privacy for mobile users (e.g., different levels of data sharing)? Does the effect of different levels of privacy on user experience and sensitivity to promotions and advertising vary? How?
Wearables and other smart devices	What are the future avenues for the growth of voice-enabled assistants? How are voice-enabled assistants likely to impact different industries, e.g., retail and hospitality? What are the ways in which mobile devices and apps can create and sustain synergies with other devices, including wearables? Are there any potential cannibalization effects? What do we know about the usage patterns across devices?
Self-driving vehicles	What are the effective ways of marketing self-driving cars? What will be the impact of driverless cars on directly- and indirectly- related markets, such as ride sharing and leasing? In what ways can retailers leverage self-driving car technologies for improved delivery?
Internet of Things (IoT)	What models can help marketers deliver timely, personalized messages and offerings to customers based on customers' stages in their shopping journeys for different products? How can marketers overcome privacy issues involved in accessing multiple devices to collect relevant data and provide useful services to customers? In what ways can marketers seamlessly develop Internet of Services (IoS) by leveraging the data from IoT?
Artificial intelligence (AI)	How do we develop AI-driven, mobile-enabled marketing systems for an array of marketing decisions ranging from product portfolio to pricing to promotion to personalization? How do we develop new models of shopping behavior of AI- and mobile-assisted shoppers and consumers? How can we apply neuroscience-based AI models that leverage mobile data to develop a deep understanding of consumer psychology?

Note: Reprinted with permission from Narang and Shankar (2019b).

Mobile Marketing Practice

The cumulative research insights and the emerging trends in mobile usage and shopping offer important implications for mobile marketing practice. Mobile marketers should consider adopting the following broad recommendations.

View mobile as an integral part of the overall marketing strategy, not just a component of the media/communication mix: Mobile is fast becoming the primary digital tool for marketers, occupying the center stage in the customer's lives and lifestyles. Given the wide range of spillover effects of any mobile intervention on customer's offline decisions, as well as the impact of offline trajectories on their mobile usage, marketers should view mobile more holistically and plan for spillovers across channels. For example, many brands, including Gillette, have adopted the direct-to-customer model to compete directly with online competition (Forbes 2016). In doing so, they have embraced mobile marketing. For example, the Gillette razor on-demand program allows customers to reorder blades simply through a text message.

Re-evaluate targeting strategy. A mobile device allows marketers to target in several ways based on locational, temporal and contextual factors. Marketing managers should carefully evaluate the different approaches. Rather than viewing the targeting decision as a one-time, piecemeal decision, managers should also consider the long-term effects of targeting and incorporate shopper learning and heterogeneity in their decisions. The combined returns from the two different types of targeting on shoppers as well as those from retargeting can be very different from a one-time, one-way targeting approach. Managers must develop creative ways of leveraging the new tools at their disposal based on data and artificial intelligence to learn from shoppers' repeated responses to different combinations of targeting approaches.

Personalize, personalize, personalize. Mobile offers unique opportunities to directly target and reach shopper on their devices. Marketers should leverage this personalization potential by designing mobile strategy that understands customer needs, perspectives, and contexts by leveraging holistic data rather than narrow data. For example, Hilton hotels allows its guests to personalize their stay and offers them digital key and check-in room selection (Forbes 2016). If personalization at the individual level is not possible in certain channels, managers should still leverage differences among subgroups of shoppers and identify shopper segments based on granular purchase and mobile device usage behaviors.

Get customers to opt-in and be transparent with customer data usage. In a data-sensitive age, the majority of customers do not appreciate unsolicited messages on their mobile devices, even if these messages are highly personalized and relevant. Respect for customer data and privacy preferences requires that firms get users to opt-in to receive mobile marketing communications. Highlighting the benefits of opting-in as well as personalizing the communications if users are already using the provider's app can be effective strategies. Firms can adopt effective strategies to get users to opt into their specific permissions by highlighting the value proposition to them before the default prompt to enable permissions in apps. For example, NHL has a full-screen 'soft-prompt' for its app users describing the benefits of notification and location enablement such as access to real-time news and game scores.

Experiment wisely. It is easy to get carried away by the hype of a new wave of innovations in wearables, augmented reality, and smart devices. Marketers need to create a unique mix and meet their customers wherever they are. Before making huge investments, learning about what customers value (e.g., do they value the convenience of voice assistants, or the control over their privacy if there are trade-offs?) can result in mutual win-wins. Again, personalization based on what kind of trade-offs each customer may be making, as learned through their revealed preferences can offer marketers a unique competitive edge.

Conclusion

Mobile marketing has gained prime importance for consumers and managers in recent years. In the omnichannel environment, managers are constantly seeking to integrate customers' experience seamlessly across devices and channels. Mobile devices, due to their personability and location-specificity, offer tremendous opportunity for engaging customers at the right time and right place. However, the full potential of mobile marketing in the Mobile 2.0 era has not been realized in research and in practice. Several trends have surfaced in areas, such as mobile app monetization, augmented reality, data and privacy, wearable devices, self-driving vehicles, the Internet of Things, and artificial intelligence. These trends present both challenges and opportunities for marketers, consumers, and society, warranting deeper investigation by marketing researchers.

In this chapter, we presented the conceptual underpinnings of mobile marketing and built on previous reviews of mobile marketing research. We offered an overview of recent advances in the mobile marketing literature by discussing three broad themes, namely, targeted mobile promotions, personalized mobile advertising, and mobile-led cross-channel effects. Within each research stream, we discussed the impact of mobile marketing on relevant customer, firm, and societal outcomes. Finally, we outlined emerging trends in practice for each of these themes and delineate the future research opportunities in mobile marketing. Our synthesis offers some useful insights, several directions for future research, and actionable implications for mobile marketing executives.

CHAPTER III

MOBILE APP INTRODUCTION AND ONLINE AND OFFLINE PURCHASES AND PRODUCT RETURNS*

In recent years, the penetration of mobile devices has reached unprecedented levels. In 2018, 77% of the United States (U.S.) population owned a smartphone (Pew Research Center 2018). Mobile *apps* are increasingly dominating mobile device use. Mobile apps account for 87% of mobile usage at 2.5 daily hours for an average adult in the U.S. (eMarketer 2017), which constitutes the bulk of digital media time (comScore 2016). Retail apps are among the fastest growing app categories (Retail Dive 2018) and the average U.S. adult has at least four retail apps on their phone (CNBC 2018b). For example, nearly 70% of Walgreens' customers engage through mobile devices and more than 50% of app adopters use the app while shopping in stores (Hyken 2017). Furthermore, about 20% of all Starbucks transactions in stores originate from its "order and pay" app (Forbes 2015).

The purpose of this chapter is to quantify the relationship between a retailer's branded mobile app launch and its online and offline purchases and product returns using data from a large-scale U.S. retailer of video games, electronics, and wireless services. In particular, we evaluate the changes in online and offline shopping behavior for app adopters relative to non-adopters, and examine the app use patterns associated

---

* Reprinted by permission, Unnati Narang and Venky Shankar, Mobile App Introduction and Online and Offline Purchases and Product Returns, *Marketing Science*, volume 38 (5), September, 2019. Copyright (2019), the Institute for Operations Research and the Management Sciences, 5521 Research Park Drive, Suite 200, Catonsville, Maryland 21228 USA.

with purchases and returns in all the channels. A mobile app may prompt shoppers to purchase more often, more items and spend more through the app. However, is app introduction and adoption linked to purchases in stores and online through the retailer's website? Furthermore, while a mobile app can prompt shoppers to purchase, does it change their product returns and the net monetary value of purchases?

The nascent literature on mobile apps is silent on how app adopters differ from non-adopters in their online and offline purchases and product returns. Two related studies focus on aggregate loyalty points or total purchases (Kim et al. 2015, Einav et al. 2014) but not on purchases by channel. Wang et al. (2015) examine the effect of a campaign promoting the use of an app on online purchases. In a cross-platform/channel study, Xu et al. (2016) examine tablet adoption's relationship with online purchases via smartphones and desktops but not through the store. Prior research has also not examined the differences between adopters and non-adopters with regard to product returns. Specifically, we address three research questions:

- Do mobile app adopters have higher or lower *frequency, quantity,* and *monetary value* of purchases and product returns across all the channels relative to non-adopters?
- How do app adopters and non-adopters differ in their purchases in existing *online* and *offline* channels?
- What potential app use patterns are associated with purchases and returns in all the channels?

We address our research questions using a unique dataset relating to app introduction by a large retailer of video games, consumer electronics and wireless services. The dataset spans 18 months before and after the introduction of the app. Teasing out the differences in app adopters' and non-adopters' shopping outcomes in the absence of randomization is a complex task. Major issues include endogeneity and self-selection of shoppers adopting the app. We mitigate the challenges posed by non-experimental variation in the data in three ways by: (a) adopting propensity score matching to identify app non-adopters similar to adopters along their observed characteristics, and comparing their shopping outcomes pre- and post- app introduction, (b) using the number of cell towers in the shoppers' zip code as a source of exogenous variation in app adoption to mitigate endogeneity due to unobservables, and (c) conducting a series of robustness tests to rule out alternative explanations.

Our results show that app adopters buy 33% more frequently, 34% more quantity, and spend 37% more than non-adopters in the period after app introduction. At the same time, they return 35% more frequently, 35% more quantity and 41% more in dollar value. Combined, app adopters offer 36% more in net monetary value (NMV) than non-adopters after app introduction. Our analyses suggest that the time, location, and features of app use provide descriptive evidence of how the app aids shopping in other channels. "App-linked" shoppers (those who make a purchase within 48 hours of app use) use the app when they are close to the store of purchase, and access the app for loyalty rewards, product details, and notifications. App adopters also purchase a more diverse set of items (including less popular products) than non-adopters. App adopters return products originally purchased in stores and through rewards more than those purchased online and without rewards.

Our research contributes to the mobile marketing and omnichannel marketing literatures in at least three important ways. First, we expand the scope of mobile marketing literature by examining mobile app adopters' and non-adopters' purchases *and* returns. Second, unlike most prior mobile marketing studies that focus on only online purchase outcomes, we consider both online *and* in-store purchases. Finally, we examine app use patterns in depth.

Our findings offer key managerial implications. We provide useful benchmarks for purchases and returns for managers to evaluate the app introduction decision. Our results also suggest that managers should plan for higher store and website visits by app adopters. Managers can predict more purchases from adopters based on their time, location, and features used in the app. Finally, managers should plan on receiving greater returns from app adopters, in particular, for products originally purchased in stores and through rewards.

Related Literature

Mobile devices influence shoppers both in- and out-of-store (Shankar et al. 2016). Mobile apps may affect purchases in three major ways. First, mobile apps can provide anytime-anywhere information benefits to shoppers. Such benefits include product- and store-related information (Danaher et al. 2015, Fong et al. 2015, Dubé et al. 2017). Second, mobile apps can offer immediate access to shopping (Shankar and Balasubramanian 2009), potentially driving impulse buying through deals. Finally, apps may serve as convenient tools and reminders for shopping.

Two streams of research, one on mobile *apps* and the other on mobile *device* adoption, are closest to our research questions. First, the sparse literature on mobile *apps* considers how app usage relates to a limited set of purchase measures. Kim et al. (2015)

study how the use of two app features influences shoppers' loyalty point accruals for an air miles reward program. They treat an app update with two new app features (not the app itself) as the intervention and find that point accruals across retailers increase. In a descriptive analysis of eBay's app usage, Einav et al. (2014) show that the app results in immediate and sustained increase in the platform's aggregate revenues. Wang et al. (2015) show that a grocery retailer's promotional campaign for a mobile app results in greater number of orders and spending. Gill et al. (2017) report similar findings for a business-to-business (B2B) app. Other papers examine related variables, such as purchase intent, website visits, and firm value (Bellman et al. 2011, Boyd et al. 2019, Xu et al. 2014).

Second, studies on mobile *device* adoption focus on the effects of mobile device on purchase frequency and monetary value. Lee et al. (2016) find that purchase frequency is higher but the monetary value of each purchase is lower for shoppers who transact more using a mobile device than the desktop. Xu et al. (2016) study the effect of tablet adoption on commerce through different devices in online retailing and conclude that tablet commerce substitutes desktop commerce but complements smartphone commerce. Overall, they find that tablet adoption enhances ecommerce revenues.

Combined, these research streams suggest that mobile apps and devices positively affect aggregate purchases but do not inform us about how mobile apps relate to outcomes in different channels, such as store and web at the shopper level. Nor do the research streams examine product returns.

Our study complements and extends these research streams as shown in Table 5.

Table 5. Selected Related Literature and Our Contribution

Paper	Focus	DV*: FP	DV: QP	DV: VP	DV: FR	DV: QR	DV: VR	Cross channel effect
Bellman et al. (2011)	Effect of app use on brand attitude and purchase intention							
Einav et al. (2014)	Analysis of eBay's mobile app adoption and platform revenues			✓				
Gill et al. (2017)	Effect of manufacturer's mobile app on business-to-business (B2B) revenues	✓	✓	✓				
Lee et al. (2016)	Effect of mobile shopping ratio (mobile vs. web) on purchases	✓		✓				
Xu et al. (2016)	Effect of tablet adoption on digital commerce via smartphones and PC devices			✓				
Our paper (2019)	Linkages between app introduction and purchase and returns across all channels	✓	✓	✓	✓	✓	✓	✓

Note: *DV refers to the key dependent variables used in the study, including frequency (FP), quantity (QP) and value (VP) of purchases and frequency (FR), quantity (QR) and value (VR) of returns. Reprinted with permission from Narang and Shankar (2019a).

First, we study the relationship between a retailers' mobile app launch (and subsequent adoption) and a variety of shopping outcomes, such as the frequency, quantity, value of purchases, and importantly, product returns. Second, we examine outcomes in existing online *and* offline channels of a retailer. Finally, we explore app use patterns that can potentially explain the relationship between app adoption and shopping in other channels.

We also draw from and extend the literature on cross-channel purchases and product returns. With regard to purchases, retail stores have a complementary relationship with the Internet channel in the long-run (Avery et al. 2012). Online sales increase after offline stores open due to gains in brand awareness in markets with low brand presence and gains in channel awareness (Bell et al. 2017, Wang and Goldfarb 2017). Furthermore, online returns decline due to enhanced product information offline (Bell et al. 2017). We apply these findings to our examination of purchases in both online and offline channels, and product returns after the app introduction.

Data and Research Setting

We collect data from a large US-based retailer of video games, consumer electronics and wireless services with 32 million customers. The gaming industry is a large $99.6 billion industry and offers a ripe setting to examine shopping behaviors. Overall, 48% of the video game purchases are physical disc format games (Venture Beat 2016).

The retailer is similar to Walmart and PetSmart, or any other brick-and-mortar chain with a relatively larger offline presence. The retailer's primary channel is its store network comprising of over 4,175 stores across the U.S. In addition, it also has an ecommerce site, but only about 5% of its sales transactions are online, although the proportion is growing fast. The retailer typically allows product returns for a full refund within 30 days of purchase. Returns take place in stores.

Our focal independent variable/intervention is app introduction. The retailer introduced its app on July 1, 2014 without any targeted campaign. Approximately 6% of shoppers adopted the app in the 18 months since launch. The app allowed shoppers to browse the retailer's product catalog, get deals and offers, order online through the mobile browsers, or locate nearby stores to buy offline (including open hours, phone numbers, and driving directions). The app allowed shoppers to add products to an in-app cart linked to a checkout button. When shoppers clicked "checkout," the app redirected them to the retailer's mobile site in the mobile browser to make an online purchase. In this way, this app is similar to several retailer apps (e.g., PetSmart). Appendix Figure A1 provides screenshots from the app.

We next describe the data used in our analysis. Mainly for security and privacy reasons, the retailer allowed us to access only a random sample of about 55,580 app adopters (from a population of about 2 million app adopters who started using the app for the first time in our data period) and 63,164 non-adopters (from the remaining 30 million shoppers).[1] The data on these shoppers include demographic information (age, gender, zip code) and loyalty program membership status. The firm provided us with two kinds of data for these shoppers, transactional data across channels and mobile app use data. From the transactional data, we have access to purchases in the firm's online and offline channels, and product returns. The online channel represents purchases made via the retailer's website, including those that come through the app checkout in the mobile browsers. In addition, we have data on product returns by the shoppers.

The mobile app data include the app features that shoppers access with timestamps. App features relate to app banner, product, stores, shopping, offers, rewards program, and notifications. We next describe these different features. App banner refers to the home screen of the app that displays offers or any other updates. Product-related features refer to product search, details, catalog, and videos that provide information about

[1] In subsequent robustness checks (Appendix Tables A1 and A2), we use an alternative sample of 50,000 app adopters and obtain similar results.

products. Similarly, store-related features refer to store-related information, such as store check-in, locations, and directions. Offers refer to a list of offers in the app or clicking on offer details to read more about a specific offer. Rewards program features include abilities to check loyalty reward points and redeem reward coupons. Finally, the notification feature refers to registering for and acting on push notifications. This feature categorization appears in Appendix Table A3.

For our research design, we collected data for a period of 18 months pre- and post-app launch. This resulted in a panel from January 2013 to December 2015. The average inter-purchase time of shoppers for the retailer's products is about 37 days. To avoid left censoring, we also collected data on the recency of purchases, that is, the day on which the shoppers made their last purchase if the last purchase was before the start of our data period.

We supplement these data with publicly available data on the number of cell towers by zip code from the U.S. Federal Communications Commission (FCC). This variable provides exogenous variation in who adopts the app, mitigating concerns about unobservable factors that might influence adopters differently from non-adopters. Higher cell tower areas are likely to get better connectivity on mobile devices and higher probability of downloading the app. Since most cell towers were constructed prior to 1990s, their location is not likely to be correlated with demand-led factors in recent years, such as population growth or demand for video games differentially in high and low cell tower zipcodes. We explain the rationale for this instrument and rule out concerns for several potential omitted variables in section 4.3.2. We show that the number of cell towers in shopper zipcode is less likely to be correlated with omitted factors that may also drive demand, e.g., store presence, store quality, income levels, etc. during our data period. The key variables, their operationalization, and descriptive statistics appear in Table 6.

Table 6. Variable Definitions and Descriptive Statistics

Variable	Operationalization	Mean	St. Dev.	Min.	Max.
Frequency of purchases	Number of purchase transactions in the time period	0.81	1.58	0	101
Quantity of purchases	Number of items bought in the time period	1.50	3.47	0	542
Value of purchases	Monetary value of purchases in the time period ($)	43.94	117.61	0	8,993.81
Frequency of returns	Number of return transactions in the time period	0.11	0.48	0	51
Quantity of returns	Number of items returned in the time period	0.15	0.73	0	120
Value of returns	Monetary value of returns in the time period ($)	3.56	29.73	0	5,150.63
Net monetary value (NMV)	(Monetary value of purchases – Monetary value of returns) ($)	40.37	107.22	-1,937	8,993.81
App Adopters (TREAT)	Dummy indicating if the shopper adopted the app (=1) or not (=0)	0.47	0.50	0	1
Time Period (POST)	Dummy indicating if the period is before (=0) or after (=1) app launch	0.50	0.50	0	1
Recency	Number of days since the shopper's last purchase	193.60	293.71	0.08	1,493.23
Age	Age of the shopper in years at the start of the data period	31.98	10.78	11.00	115
Gender	Gender of the shopper (Female=2, Male=1, Unknown=0)	0.67	0.63	0.00	2.00
Distance to nearest store	Distance in miles between the geographical centers of the shopper's and the nearest store's zip codes	3.90	7.21	0.00	574.48
Number of stores	Number of the focal retailer's stores in the shopper's zip code	0.56	0.72	0.00	4
Loyalty program level	Dummy indicating if the shopper is enrolled in the basic (=0) or professional (=1) membership category on app introduction date	0.16	0.37	0.00	1
Area population	Population of the shopper's zip code based on 2010 US census	31,437	19,090	6.00	113,916
Competitor stores	Number of competing stores in the shopper's zip code	0.29	0.52	0.00	4
No. of unique products	Number of unique SKUs (Stock Keeping Units) that the shopper buys	1.21	2.71	0.00	313
No. of unique categories	Number of unique categories that the shopper buys	0.84	1.53	0.00	29
Pct. of top 100 products	(Spending on the Top 100 products/Total spending)	0.21	0.34	0.00	1
Pct. of top 500 products	(Spending on the Top 500 products/Total spending)	0.47	0.42	0.00	1

Table 6 continued

Variable	Operationalization	Mean	St. Dev.	Min.	Max.
Cell towers	Number of cell towers in the shopper's zip code	6.99	6.93	0.00	52
Precipitation	Average precipitation level in the shopper's zip code as reported by NOAA in millimeters in June 2014	102.82	74.81	0.00	498.5
Temperature	Average air temperature of the shopper's zip code in degree Celsius as measured by NOAA in June 2014	23.53	3.57	0.96	36.38
Download speed	Percentage of the population in the shopper's county with download speeds less than 6,000kbps as reported by FCC in June 2014	0.01	0.02	0.00	0.81
Wireless access	Percentage of the population in the shopper's county with access to three or more wireless providers as reported by FCC in June 2014	0.12	0.02	0.00	1

Note: The statistics for the outcome variables (frequency, quantity and value of purchases and returns) are averaged over the 36-month period. Reprinted with permission from Narang and Shankar (2019a).

Analyses

Relationship between App Introduction and Shopping Outcomes: Descriptive Analysis

We define app adopters as shoppers who started using the app for the first time during our data period. Non-adopters are those who did not access the app even once during the study period.[2] A simple comparison of shopping outcomes shows that the average monetary value of purchases increased 19.25% ($63.60 to $75.84 per month) for app adopters, while it decreased marginally by 2.92% ($21.56 to $20.93 per month) for app non-adopters after app introduction ($p < 0.001$). However, the average monetary value of returns for app adopters also increased by 19.37% ($5.37 to $6.41 per month) compared to app non-adopters who experienced marginal decrease of 8.23% ($1.58 to $1.45 per month) in the same period ($p < 0.001$). Overall, the net monetary value increased by more than 19.23% for app adopters compared with 2.53% decline for non-adopters. Notably, both the online and offline purchases are higher for app adopters than for non-adopters by 55.77% and 21.34%, respectively ($p < 0.001$). For the matched sample of non-adopters in Table 7, the comparison shows similar direction but larger magnitude of difference in treated and *matched* control relative to differences in treated and *unmatched* controls over the same time period. Between the pre- and post- periods, the value of purchases decreased by 17.45% (from $63.10 to $52.09) for the matched control shoppers and value of returns decreased by 21.10% (from $5.26 to $4.15).

---

[2] The findings of the chapter apply to app "adopters" defined as those who download *and* use the app, not those who download but never use the app.

Table 7. Model-free Evidence: Mean Statistics

Variable	Treated pre period	Treated post period	Control pre period	Control post period	Matched controls pre period	Matched controls post period
Frequency of purchases	1.211	1.368	0.407	0.363	1.197	0.963
Quantity of purchases	2.256	2.524	0.570	0.665	2.214	1.737
Value of purchases	63.60	75.84	21.56	20.93	63.10	52.09
Frequency of returns	0.177	0.196	0.047	0.041	0.169	0.129
Quantity of returns	0.238	0.258	0.062	0.053	0.226	0.167
Value of returns	5.369	6.406	1.584	1.452	5.257	4.146
Net monetary value of purchases	58.236	69.438	19.975	19.481	57.85	47.94
Frequency of purchases - online	0.019	0.033	0.006	0.007	0.020	0.017
Quantity purchased - online	0.028	0.050	0.009	0.010	0.028	0.025
Value of purchases - online	1.41	1.997	0.495	0.425	1.571	1.085
Frequency of purchases - stores	1.192	1.335	0.400	0.356	1.177	0.946
Quantity purchased - stores	2.228	2.474	0.741	0.655	2.186	1.713
Value of purchases - stores	62.20	73.85	21.06	20.51	61.53	51.01

Notes: The pre-period (post-period) statistics are monthly averages over the 18-month period pre (post) app launch across shoppers; online purchases in the post period for the treated include purchases on the mobile site after clicking checkout in app; matched controls are non-adopters similar to adopters based on Nearest Neighbor matching using pre-period covariates. Reprinted with permission from Narang and Shankar (2019a).

Econometric Model

To estimate the effect of mobile app introduction, the ideal approach would be to compare the shopping outcomes for shoppers who adopt the mobile app to the counterfactual, that is, to outcomes when the same shoppers do not adopt the mobile app. However, because we do not observe the counterfactual (a shopper cannot be both an app adopter and a non-adopter) and because the treatment is not randomly assigned (a shopper self-selects into adopting the app), we examine non-experimental variation. Specifically, we compare the pre- and post- app introduction outcomes for app adopters and similar non-adopters. After specifying the baseline regression model, in the next section, we outline our strategy to address the endogeneity of individual app adoption decision using propensity score matching (PSM, Rosenbaum and Rubin 1983) and the Heckman selection model. We rule out several competing explanations for our results in the robustness checks. The complete list of analyses appears in Table 8.

Table 8. Overview of Analyses

Section	Analysis	Objective	Key insight/Conclusion
4.3	Propensity Score Matched and Heckman Two-Stage Linear Regression	Quantifying the relationship between app introduction and shopping outcomes	App introduction is related with higher frequency, quantity, and monetary value of online and offline purchases, returns, and net monetary value of purchases.
5.2 and Appendix	<u>Robustness Checks</u> (a) Alternative models: Diff-in-diff without matching, and Poisson Diff-in-Diff (b) App adoption date as cut-off (c) App novelty (d) Alternative matching (e) Outliers (f) Shopper heterogeneity in deal proneness (g) Alternative sample	Ruling out alternative explanations for the results	App introduction effects are robust to alternative specifications and explanations, such as other adoption measures, time periods, app novelty, customer deal-proneness, and outliers.
6	<u>App Use Patterns</u> Time, geography and features used	Identifying app use patterns associated with shopper purchases and returns	Increased purchases seem to be associated with app use time, geography and features; adopters buy diverse and less popular products; returns are higher for adopters' in-store and rewards-based purchases.

Note: Reprinted with permission from Narang and Shankar (2019a).

We compare the change in outcomes for the app adopters 18 months before and 18 months after the retailer introduces the app to the change in outcomes for the non-adopters over the same time period. Formally, we specify our difference-in-differences linear regression model as:

$$(1) \quad Y_{it} = \alpha_0 + \alpha_1 TREAT_i + \alpha_2 TREAT_i * POST_t + \tau_t + \vartheta_{it}$$

where i is individual, t is month, Y is the outcome variable (frequency, quantity and monetary value of purchases and returns), $TREAT$ is a dummy variable denoting app adoption (1 if shopper i is an app adopter and 0 otherwise) to account for time-invariant unobserved differences in adopters and non-adopters, $POST$ is a dummy variable denoting the period (1 for the period after the app has been introduced and 0 otherwise), $\boldsymbol{\alpha}$ is a coefficient vector, τ represents month fixed effects and ϑ is an error term. The coefficient of TREAT x POST represents the change in app adopters' outcomes relative to non-adopters' outcomes between the pre- and post- periods.

The challenge in identifying the model specified in equation (1) is that app adoption is endogenous. In other words, shoppers choose to adopt an app in anticipation of future purchase behaviors and because of their inherent characteristics, both observed and unobserved (e.g., preference for the retailer). The selection challenges are posed by non-experimental variation in our data similar to quasi-experiments (Tirunillai and Tellis 2017). Next, we describe empirical approaches for mitigating concerns for selection on observables and unobservables.

Endogeneity and Self-Selection

The first availability of the retailer's mobile app represents a shock in our data period. The timing of the introduction of the app is exogenous to the shoppers, but the date on which an individual shopper adopts and starts using an app may be endogenous.

Therefore, we use the app introduction date as the cut-off for defining the pre- and post-periods. In this way, our empirical strategy conservatively assumes that the benefits of the app start accruing right from the day of app introduction for all app adopters, irrespective of their self-selected time of app adoption (Manchanda et al. 2015). An alternative empirical strategy for estimating the effect of the app would be to treat the app adoption date for each adopter as the cut-off date for the intervention. This approach does not rule out endogenous app adoption timing because app adopters may decide to adopt the app in anticipation of increased future purchases. As a result, our main analysis reports the results from the first approach. However, as a robustness check, we estimate a model using the second approach that leverages individual adoption dates (see Section 5.2.2).

Arguably, one concern with using the app introduction date is the question of what exactly are we measuring. Is it the effect of app introduction or adoption by shoppers? Our context is similar to Manchanda et al. (2015). They study the launch of an online community that users join endogenously at different points in time. They use the launch of the community as the intervention and leverage the period before- and after- the community launch as their pre- and post- intervention periods. However, they attribute the coefficients of "treatment effect" of launching the community to users' act of joining the community.

In our setting, while we use the launch date to quantify the shift pre- and post- app introduction, the differences in shopping outcomes come from those who adopt the app (and not from their purchases even before they have adopted). We triangulate this through several analyses and falsification checks (see section 4.3.2). In this sense, we measure the relationship between app introduction and shopping outcomes for the shoppers adopting the app. Therefore, our estimates are more representative of a local average treatment effect (LATE) than an average treatment effect (ATE) for the population.

In the absence of randomization on who adopts the app, we use several careful econometric techniques to examine the effects of app introduction. First, we use propensity score matching based on observed individual covariates. We conduct several tests to inspect the quality of the matches. Second, we estimate a two-stage Heckman selection model using the exogenous variation in the number of cell towers in the shopper's zip code to mitigate selection on unobservables. Third, we carry out a set of additional falsification analyses. We next discuss our approaches for selection on observables and unobservables.

Selection on Observables: Propensity Score Matching

Propensity score matching allows us to match app adopters and non-adopters on observed demographic and behavioral covariates, while tackling the curse of dimensionality (Rubin 2008, Guo and Fraser 2014). We begin by calculating each shopper's propensity score, defined as the shopper's probability of adopting the app. We do this using a binomial logit model (see Appendix Table A4). Next, we identify non-adopters similar to adopters based on the estimated propensity scores to create a control group. This approach is consistent with Rosenbaum and Rubin's (1983) in creating a control group similar to the treated group in the distribution of observed covariates. We match each app adopter to a non-adopter based on the 1:1 Nearest Neighbor algorithm with replacement. Formally, if $P(X_i)$ is shopper *i*'s propensity score, the treated shopper *i* is matched to the control individual *j*, where *j* is min $||P(X_i) - P(X_j)||$ to create matched pairs closest to each other (Wangenheim and Bayón 2007, Huang et al. 2012).

What factors explain the decision to adopt the app? Consistent with extant research (Hung et al. 2003; Kim et al. 2015), we model app adoption as dependent on individual shopper demographics (e.g., age, gender), behavioral measures (e.g., pre-period spending and returns by month, recency, and frequency of purchase, online purchases) and other related measures (e.g., distance to the nearest store, number of stores in the shopper's zip code, presence of competitor stores, loyalty program membership level on app introduction day) that are likely to influence shoppers.

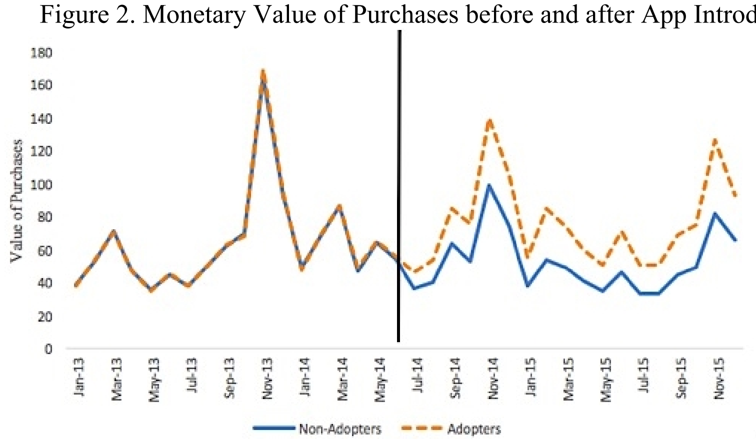
(2) Propensity/Probability of App Adoption $A_i = \frac{\exp(U_i)}{1 + \exp(U_i)}$

(3) $U_i = \gamma + \delta D_i + \varepsilon_i$

where *i* is shopper, *U* is the utility from app adoption, *D* is a vector of covariates, (γ, δ) is a coefficient vector, and ε is an error term, distributed as double exponential (Wooldridge 2002). To test the goodness of our propensity score matches, we conduct a set of statistical analyses, including the Standardized Bias Reduction test (Rosenbaum 2005).

Matching improves the percentage balance of propensity scores by 100%, making the matched treated and control groups comparable (see Appendix Table A5). This is important evidence that matching results in a valid control group. Another evidence is visual verification that the pre-period purchase trends for the treated and control groups are parallel (Figure 2).[3] Appendix Table A6 presents the model estimates from the same comparison shown in Figure 2, based on the model in equation (1) with treatment interacted with each monthly dummy. This check together with the insensitivity of results to control variables (discussed in Section 5.2.1), assures us that we can reasonably mitigate selection concerns along observables.

Figure 2. Monetary Value of Purchases before and after App Introduction



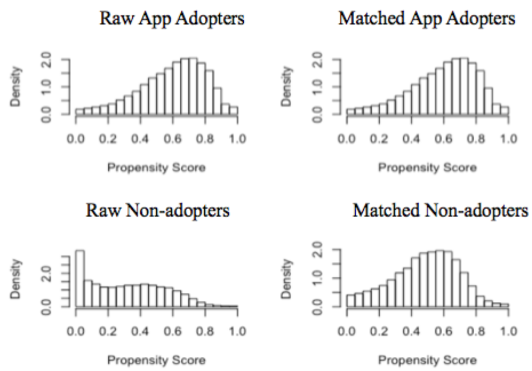
Notes: The solid black vertical line represents the app introduction time. In the pre period, the values of purchases are nearly identical for the adopters and non-adopters, hence appearing as a single line. The sharp peak in November 2013 reflects the launch of Xbox One console, which sold over 2 million units worldwide within the first month. Reprinted with permission from Narang and Shankar (2019a).

Histograms showing the distribution of propensity scores for the treated and control groups before and after matching appear in Figure 3. The distributions visually look more similar after matching relative to before matching.

---

[3] We report the pre-period trends for the unmatched samples (raw data) in Appendix Figure A2. Together with Figure 2, it demonstrates that matching improves the pre-period trends, mitigating concerns for unobservables.

Figure 3. Distribution of Propensity Scores Pre- and Post-Matching



Note: Raw adopters' and non-adopters' plots represent pre-matching propensity scores. Reprinted with permission from Narang and Shankar (2019a).

Matching estimates and inspecting common trends in our data ameliorate the self-selection concern to some extent as shown in Figure 3. Nevertheless, we also carry out several other analyses to further reduce concerns for unobserved omitted variables (e.g., unobserved preference for the firm).

Selection on Unobservables: Heckman Correction and Other Analyses

To more formally account for the non-randomness of app adoption due to unobserved factors, in this section, we carry out two types of analyses. In the first set of analyses, we use: (a) the two-stage Heckman correction procedure (Heckman 1979) and (b) an alternative control group based on future treated shoppers because they are more similar to app adopters than are non-adopters. In the second set of analyses, we exploit falsification tests to check that the estimates are not spurious by comparing app adopters' outcomes relative to non-adopters' in the periods after app introduction but prior to the adopters' specific adoption dates, consistent with Goldfarb and Tucker (2014).

We first describe the Heckman selection model. In the first stage, we model the choice to adopt the app using a probit model. App adoption is an endogenous variable in our outcome model since shoppers self-select into adopting the retailer's app. To identify the model, we require exogenous variation in app adoption. A potential source of such variation is the number of cell towers in the shopper's zip code. Presumably, a higher number of cell towers may be associated with better wireless connectivity on the shoppers' mobile devices. This is likely to influence app adoption. To meet the exclusion restriction, the number of cell towers should not be correlated with the error term ϑ_{it}. In other words, it should not drive demand through variables other than app adoption. Consistent with Bronnenberg et al. (2012) and Lundborg et al. (2017), we carefully consider and argue against several potential omitted variables that may be correlated with the number of cell towers.

First, in our context, high population growth areas or areas that experienced recent population surges could possibly endogenously drive the number of cell towers. If so, in such locations, the number of cell towers may be correlated with the demand-related factors. We analyze relevant data on population from recent years and the number of cell towers for the zip codes in our data. We examined the correlations between the number of cell towers constructed and population in a zip code during 2010-2015. The correlation is low and positive (0.072). Similarly, the correlation between the number of towers and the change in population between 2010 and the start year of our data, namely, 2013, is low and negative (-0.004). Furthermore, we find that 90% of the cell towers in our data set were constructed prior to 2010 with 50% developed prior to 1990, before the birth of commercial Internet. Thus, the changes in population and the resulting demand effects during our period of data would have less likely contributed to the variation we observe in cell towers across zip codes due to population growth. Furthermore, 42% of the cities in the data show stagnant or negative growth in population, ameliorating a concern that the emergence of cell towers in these already built-up cities could have been in response to population growth in recent years. In areas where the locations of recently constructed cell towers map closely with population density, our exclusion restriction could fail.[4] However, in our nationwide data, this is the case for only 0.6% of the locations.

Second, a possible alternative variable through which the number of cell towers can directly affect shopping outcomes is store availability. However, the correlation between the number of towers and the number of stores in a zip code is low at 0.18. During 2010-2015, the correlation between the number of new stores opened and the number of cell towers is even lower at 0.012. Nonetheless, we also control for population, number of stores, and the distance of shopper to the nearest store in our matching algorithm. Thus, both the app adopters and the matched non-adopters are similar in these variables besides the application of Heckman selection correction.

Third, another potential alternative variable through which the number of cell towers can directly affect shopping outcomes is store performance. Perhaps the retailer's stores in high cell tower regions perform better than stores in low cell tower regions because

[4] The limits and scope of our exclusion restriction is similar in spirit to Jacobi and Sovinsky's (2016). They estimate marijuana demand using living in a city and temperature as exclusionary variables.

they might have superior staff and product assortment. To examine this possibility, we collect data on two commonly used measures of store performance, sales and sales per square foot (Gomez et al. 2004). We estimate the differences in these store performance measures across high and low cell tower zip codes based on median split and find that there is no significant difference in the average store sales ($p > 0.10$) and average store sales per square foot ($p > 0.10$) during the data period.

Fourth, the number of cell towers could also affect shopping outcomes if high cell tower regions are those with stores offering better product assortment. We also confirm from the managers of the retail chain that they do not strategically vary product assortments across stores, as well as between online and offline channels. This fact mitigates the concern that product assortment might vary between stores in high cell tower areas and stores in low cell tower areas.

Fifth, the number of cell towers may be endogenous because they might have been constructed in areas needing greater wireless connectivity. We find that most of the cell towers were built prior to 1990. However, even if the number of cell towers is highly correlated with the need for wireless connectivity, this demand is less likely to come from console buyers who primarily play games offline in the console and do not depend on connectivity to the Internet. This focal retailer primarily sells console games. Furthermore, online game playing is a recent phenomenon, is still a small fraction of console game playing time and occurred after most cell towers were built. About 70% console buyers play games that are offline and their experience is not tied to better connectivity (Statista 2015). Anyone can purchase a game and play on their home console without being connected to the Internet. Thus, in terms of scope and generalizability, our results are best interpreted in the context of retailing of an offline product, e.g., electronics, consoles and offline games, rather than online and mobile games that require Internet connection.

Sixth, the location of cell towers may be correlated with economic variables, such as income and employment. To examine this possibility, we analyzed the correlations by zip code. The correlation between the number of towers and the mean (median) household income in 2013 inflation-adjusted dollars is low at -0.13 (-0.11). Similarly, the correlation between the number of towers and employment rate (measured as the % population 16 years and over in labor force) is low at 0.03.[5]

Next, we estimate a probit model to examine the strength of the first-stage effects. The exclusion restriction results in the following first-stage selection equation for modeling A_i, the probability of shopper *i* adopting the app:

$$(4)\quad \Pr\left(A_i = 1 \middle| CELLTOWER_i, Q_i, \epsilon_i\right) = \Phi(\partial_0 CELLTOWER_i + \partial_1 Q_i + \epsilon_i)$$

where CELLTOWER is the number of cell towers in the shopper's zip code, Q is a vector of other covariates similar to those used in propensity score matching, $\partial$is a coefficient vector, and ϵ is an error term. We compute the inverse Mills ratio from this probit regression. We then augment the regression model in equation (1) by including the inverse Mills ratio as an additional covariate in the second stage. Because the number

[5] Despite the thorough evaluation of the exclusion restriction and other falsification checks, we cannot completely rule out the possibility that there could be other unobserved factors not addressed by the number of cell towers variable (Goldfarb and Tucker 2011). If there is any such unobserved factor, we need to be cautious in interpreting our estimates as causal (Jacobi and Sovinsky 2016).

of cell towers vary at the zip level, we cluster standard errors by zip code to allow serial correlations and within-zip code correlations among shoppers (Bertrand et al. 2004).

Another approach we adopt for selection on unobservables is to use an alternative control group. In this approach, we treat future app adopters as a control group. Specifically, we consider a shopper who adopted the app in the nine months between April 2015 and December 2015 (both months included) to be in the control group while we view those who adopted the app in the first nine months of app launch in July 2014 to be in the treatment group. We compare the outcomes of these groups of shoppers nine months before and after app launch. This analysis is in the same spirit as the falsification analysis of Manchanda et al. (2015) and Goldfarb and Tucker (2011).

Finally, we report a falsification test to formally check whether app introduction has a strong relationship with shopping outcomes right from the time of launch or from when it is adopted by the shoppers. This test will help mitigate any concern about an unobserved event around app introduction that may be leading to spurious effects. Introduction by itself, in the absence of adoption, should not be linked to significant upticks in purchases despite possible positive spillover effect of the news of app release. The way we conduct this falsification check is to first create different cohorts of app adopters based on adoption date. We then define the "post" period observations as those from the months after app introduction but before the cohort's app adoption date. Thus, adopters of August 19, 2015 will have their post-period defined as July 2014 through July 2015. Their pre period will be the same number of months before app launch, i.e., June 2013 through June 2014. We then estimate a regression model for these periods for the treated cohort with each adopting shopper's matched non-adopter. While this test is not an exact replication of the full model due to reduced number of post-period observations, it lends some confidence in the estimates by showing the lack of significant effects in the months before adoption.

Results and Robustness Checks

Table 9 shows the estimates from the first-stage self-selection probit model.

Table 9. First-Stage Probit Model Results

Variable	Coeff. (Std. Err.)
Cell towers	0.0022***(0.0006)
Precipitation	0.0003***(0.0001)
Temperature	0.0036**(0.0012)
Download speeds (less than 6,000k)	-0.0111(0.0168)
Wireless access	0.0075(0.0184)
Age	-0.0142***(0.0004)
Gender	0.1566***(0.0066)
Recency	-0.0018***(0.0000)
Loyalty program level	-0.1902***(0.0106)
No. of stores in zip code	-0.0168*(0.0067)
Area population	0.0000***(0.0000)
Distance to nearest store	-0.0004(0.0007)
No. of competitor stores	0.0201*(0.0083)
Past frequency of purchases	0.0003***(0.0000)
Past quantity of purchases	-0.0003***(0.0001)
Past monetary value of purchases	0.0001***(0.0000)
Past frequency of returns	-0.0024***(0.0010)
Past quantity of returns	-0.0019(0.0062)
Past monetary value of returns	0.0225***(0.0016)
Past frequency of purchases – online	-0.0001(0.0039)
Past quantity of purchases – online	0.0406***(0.0151)
Past monetary value of purchases – online	0.0145*(0.0074)
Past quantity of purchases per order	-0.0003***(0.0001)
Intercept	-0.0084(0.0363)

Notes: The number of cell towers in the shopper's zip code provides the exogenous variation in app adoption; AIC = 127,565.6; BIC = 127,798 *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Reprinted with permission from Narang and Shankar (2019a).

A higher number of cell towers is associated with a higher adoption probability in the first stage ($p < 0.001$).

We include the inverse Mills ratio (IMR) obtained from the first stage in the second stage outcome model for the propensity score matched sample. The results for this model appear in Table 10. IMR is significant ($p < 0.001$). The coefficient of IMR is the fraction of the covariance between the decision to adopt the app and shopping outcome relative to the variation in decision to adopt the app. The coefficient of IMR is indeed negative (similar to Gill et al. 2017). This sign indicates that the selection correction term, resulting from omitted factors affecting shoppers, adjusts shopping outcomes downward.

The results in Table 10 show a positive and significant relationship between app introduction on the frequency, quantity and monetary value of purchases and returns ($p < 0.001$). These estimates are more representative of a local average treatment effect (LATE) than an average treatment effect (ATE) for the population. They are particularly relevant for app adopters whose decision to adopt the app is affected by the number of cell towers in the zip code, consistent with Sudhir and Talukdar (2015). App adopters spend \$23.26 more than non-adopters after app introduction and engage in a higher frequency ($\alpha_2 = 0.39$, $p < 0.001$) and quantity of purchases ($\alpha_2 = 0.75$, $p < 0.001$). Interestingly, relative to non-adopters, app adopters also return \$2.15 more of products and engage in a greater frequency ($\alpha_2 = 0.06$, $p < 0.001$) and quantity of returns ($\alpha_2 =$

0.08, p < 0.001). Overall, app adopters' net monetary value of purchases is 36.49% higher than matched pre-period non-adopters'.[6]

Table 10. App Introduction and Aggregate Shopping Outcomes

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
(App Adopters X Post App Introduction)	0.391*** (0.014)	0.745*** (0.027)	23.256*** (0.75)	0.059*** (0.004)	0.080*** (0.008)	2.149*** (0.217)	21.107*** (0.671)
App Adopters	0.018 (0.022)	0.051 (0.041)	0.746 (1.165)	0.009 (0.008)	0.013 (0.013)	0.134 (0.374)	0.612 (0.986)
IMR	-2.446*** (0.044)	-4.504*** (0.082)	-122.22*** (2.082)	-0.380*** (0.016)	-0.508*** (0.025)	-11.232*** (0.623)	-110.99*** (1.794)
Intercept	7.148*** (0.124)	13.123*** (0.23)	350.28*** (5.945)	1.094*** (0.045)	1.456*** (0.071)	32.731*** (1.811)	317.55*** (5.115)
Number of Observations	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760
R squared	0.156	0.117	0.098	0.047	0.038	0.012	0.099

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value. IMR = inverse Mills ratio from the selection model. Reprinted with permission from Narang and Shankar (2019a).

Furthermore, both the online and offline purchases tend to be higher for app adopters relative to non-adopters (Table 11). The absolute effects are much larger for in-store purchases ($22.18 additional amount) than online purchases ($1.07 additional amount). However, in percentage terms, the increase in online purchases for app adopters relative to matched non-adopters is 68% compared to about 36% for corresponding in-store purchases (p < 0.001).

[6] We calculate the percentage change as the treatment effect relative to matched pre-period control group's level of the variable reported in Table 3. For example, for NMV, it is (21.11/57.85) or 36.49%.

Table 11. App Introduction and Purchases by Channel

Variable	(1) Offline Frequency of purchases	(1) Offline Quantity of purchases	(1) Offline Value of purchases	(2) Online Frequency of purchases	(2) Online Quantity of purchases	(2) Online Value of purchases
(App Adopters X Post App Introduction)	0.375*** (0.014)	0.719*** (0.027)	22.183*** (0.737)	0.017*** (0.001)	0.026*** (0.002)	1.074*** (0.123)
App Adopters	0.019 (0.021)	0.051 (0.041)	0.902 (1.143)	0.000 (0.001)	0.000 (0.002)	-0.157 (0.130)
IMR	-2.411*** (0.044)	-4.454*** (0.081)	-119.91*** (2.049)	-0.035*** (0.001)	-0.05*** (0.003)	-2.306*** (0.122)
Intercept	7.042*** (0.123)	12.974*** (0.229)	343.72*** (5.855)	0.106*** (0.004)	0.149*** (0.008)	6.56*** (0.343)
Number of Observations	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760
R squared	0.050	0.038	0.051	0.003	0.002	0.003

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. IMR = inverse Mills ratio from the selection model. Reprinted with permission from Narang and Shankar (2019a).

Our findings show a significant change in shopper behavior in other channels. We find that app adopters buy 33% more frequently, 34% more quantity and spend 37% more than non-adopters after app introduction. At the same time, they return 35% more frequently, 35% more quantity and by 41% more in dollar value compared with non-adopters after app introduction. Overall, the app is associated with a 36% increase in net monetary value of purchases across the channels. Both the online and offline purchases are higher for app adopters.

Similarly, the main model with an alternative control group based on future adopters produces robust estimates. These results appear in Table 12.[7] In this method, we use

[7] The results from this model are similar for online and offline purchases as reported in Appendix Table A7.

matched future app adopters from April-December 2015 as comparable non-adopters for app adopters from July 2014-March 2015 and compare their outcomes pre- and post- nine months of app introduction. The results from this model are substantively similar.

Table 12. Alternative Model with Future App Adopters as Control

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
(App Adopters X Post App	0.262***	0.482***	15.534***	0.045***	0.063***	1.789***	13.745***
Introduction)	(0.014)	(0.032)	(1.196)	(0.005)	(0.007)	(0.271)	(1.082)
App	0.007	0.015	0.345	0.000	-0.003	0.054	0.291
Adopters	(0.022)	(0.047)	(1.479)	(0.007)	(0.011)	(0.294)	(1.334)
IMR	-3.248***	-5.96***	-173.679***	-0.518***	-0.696***	-16.03***	-157.65***
	(0.078)	(0.148)	(4.15)	(0.017)	(0.025)	(0.633)	(3.731)
Intercept	9.745***	17.911***	508.894***	1.55***	2.096***	44.981***	463.913***
	(0.205)	(0.395)	(11.007)	(0.048)	(0.07)	(1.713)	(9.914)
Number of Observations	1,388,772	1,388,772	1,388,772	1,388,772	1,388,772	1,388,772	1,388,772
R squared	0.058	0.041	0.055	0.014	0.011	0.004	0.059

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; in this method, matched future app adopters from April-December 2015 are used as controls for app adopters from July 2014-March 2015; their outcomes pre- and post- 9 months are compared; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$; NMV = Net Monetary Value. IMR = inverse Mills ratio from the selection model. Reprinted with permission from Narang and Shankar (2019a).

Finally, we consider the pre-adoption period for the app adopters as their post period relative to the pre-launch period by adoption date-based cohorts. The estimates for nine such cohorts (randomly drawn from adoption dates between April-December 2015) appear in Table 13. We report the coefficient of TREAT x POST variable for the adopters on each of the dates relative to their matched non-adopters. The "post" period observations are those from the months after app introduction but before app adoption.

From Table 13 most coefficients are insignificant. Even so, for the cohorts with significant coefficients in Table 13, we do not spot any systematic pattern. Thus, the availability of the app does not appear to be related to shopper behavior in the absence of adoption. The results in Table 13 are for the shorter-term period relative to the main model period of 18 months. Thus, these estimates are best viewed as depicting short-term relationships.

Table 13. Falsification Test with Post-Period defined as Post App Introduction but Pre App Adoption

Cohort	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
4/16/15	0.139	0.219	-2.511	0.064	0.068	6.218	-8.729
	(0.14)	(0.287)	(12.827)	(0.052)	(0.07)	(3.443)	(11.557)
5/10/15	0.287	0.561	22.386*	0.068	0.088	3.993	18.394
	(0.147)	(0.316)	(10.948)	(0.042)	(0.059)	(2.407)	(9.957)
6/28/15	0.038	-0.179	-0.249	-0.03	-0.051	-1.28	1.031
	(0.11)	(0.229)	(10.03)	(0.036)	(0.064)	(2.629)	(9.462)
7/18/15	0.159	0.058	5.235	0.023	0.025	0.398	4.837
	(0.123)	(0.248)	(9.485)	(0.039)	(0.055)	(2.547)	(8.434)
8/19/15	0.168	0.164	-10.622	0.085*	0.089	1.296	-11.918
	(0.119)	(0.252)	(9.674)	(0.04)	(0.052)	(1.576)	(9.464)
9/22/15	0.337*	0.718*	19.026	-0.013	-0.002	-2.928	21.953*
	(0.145)	(0.351)	(12.749)	(0.04)	(0.059)	(2.83)	(11.109)
10/21/15	-0.127	-0.062	5.929	-0.06	-0.086	-2.602	8.531
	(0.156)	(0.291)	(10.365)	(0.039)	(0.061)	(1.596)	(9.964)
11/8/15	0.033	0.039	-1.064	0.008	0.011	-1.076	0.013
	(0.139)	(0.286)	(11.122)	(0.03)	(0.04)	(1.62)	(10.746)
12/15/15	0.056	-0.024	-12.333	0.058	0.084	-0.7	-11.633
	(0.112)	(0.24)	(11.605)	(0.052)	(0.062)	(3.237)	(10.541)

Notes: Coefficients of TREAT x POST are reported in this table for cohorts based on their date of app adoption for randomly selected dates between April and December 2015; robust standard errors clustered by zip code are in parentheses; month fixed effects are included; post period is defined as post app introduction but pre app adoption; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. NMV = Net Monetary Value. Reprinted with permission from Narang and Shankar (2019a).

Checking Robustness of Results and Ruling out Alternative Explanations

We perform several robustness checks and tests to rule out alternative explanations for the linkages we find between app introduction and purchases and returns.

Alternative model specifications: In addition to our preferred model with propensity-score matched linear regression and Heckman selection correction, we also estimate models without matching and using Poisson count data models for frequency and quantity variables. The results from these models replicate the findings from Tables 10 and 11 and appear in the Appendix Tables A8-A11. Coefficients of the treatment effect from Tables A8 and A9 represent unconditional changes in outcomes due to the app, without conditioning on covariates via propensity scores. Tables A10 and A11 report the results from Poisson count data models. These results are substantively similar to those in Tables 10 and 11 for the matched samples. The insensitivity of the results to control variables suggests that the effect of unobservables relative to these observed covariates would have to be very large to change our results significantly (Altonji et al. 2005).

Alternative cut-off as app adoption: We estimated an alternative model with a more nuanced pre- and post- period cut-off date based on individual app adoption. In our main model, we compared app adopters and non-adopters along their outcomes before and after app introduction. In the alternative model, we create a new setup where we treat the app adoption date for each treated shopper as the cut-off date for the intervention. In this model, $POST_t$ in the regression from our estimation equation (1) becomes $POST_{it}$, a binary variable that indicates the post app adoption period for each treated shopper and the matched counterpart. For example, if a shopper first adopted the app in September

2014, the indicator value is 1 for subsequent months, else 0. The matched shopper for this adopter will share the same value. Thus, the comparison time window is the same for each matched pair (Xu et al. 2016). The results appear in Appendix Tables A12 and A13. The effects are robust and consistent with previous findings.

App novelty effect: An alternative explanation for the increased net monetary value of purchases after app adoption could be the novelty of the app. It is possible that the app triggers a heightened shopping response only due to a temporary novelty effect around its launch that fades after a few months. To test this explanation, we re-estimate our models using different windows of time, including three, six, nine, and 12 months pre- and post- introduction for the matched sample of adopters and non-adopters. The effects of the app introduction persist in these varying windows of time (see Appendix Tables A14 and A15).

Alternative matching methods: Our main analysis relies on the commonly used 1:1 nearest neighbor matching algorithm with replacement. In addition, we use the Mahalanobis metric, the nearest neighbor without replacement and a refined caliper matching approach by defining the bandwidth within which to identify matched control units (Silverman 1986). We test using a bandwidth of 0.27 and a tighter 0.05 times the standard deviation of the propensity scores. Appendix Tables A16 and A17 report the results for these alternative matched samples. We find that these estimates are consistent with those from our proposed method.

Outliers: Another possible explanation for the effects could be outliers, such as the top spenders (and not the average shopper). We test this possible explanation by first removing the top spenders from our sample (the mean plus three standard deviations of spending by shoppers in the pre period) and then carrying out difference-in-differences methods as done earlier. Our estimates are robust to outliers as shown in Appendix Tables A18 and A19.

Shopper heterogeneity: While we match shoppers on a broad set of covariates, an untested alternative explanation is that deal-prone shoppers rather than app usage largely drive the effects. In general, the retailer did not send any unique offers to mobile app adopters that the non-adopters did not receive, or vice-versa. Yet, to rule out the possibility that actual *redemption* or *use* of offers in the pre-period could influence the two groups differently, we repeat the analyses after removing deal-prone shoppers. Appendix Tables A20 and A21 report the estimates for the sample that did not use deals in the pre period. The results are substantively similar.

Alternative sample: While we analyze a sample of 55,580 app adopters (about 3% of all adopters) in our main analysis, we collect additional data on another random sample of 50,000 app adopters. We repeat the analyses for this sample using a propensity score matched linear regression with Heckman section correction from our main model. We find consistent results for this alternative sample. Appendix Tables A1 and A2 report these estimates.

App Use Patterns: An Exploratory Examination

Our results show that app adoption and not its availability is primarily associated with shopping outcomes (see Table 13). In this section, we examine app adopters' app *use* and offer a descriptive analysis of the linkages between shoppers' app use patterns and purchase or return transactions.

Specifically, we examine three elements of app use, namely, time, geography, and features, in the six months following the launch of the app.[8] We identify and report differences in these usage aspects in these six months between app adopters who subsequently make a purchase ("app-linked" shoppers; N= 21,092) within 48 hours after app use and those who do not ("non-app-linked" shoppers; N = 8,934). We chose 48 hours based on the distribution of the time gap between app use and purchase. We examined the distribution of time gap between app use and purchase time (within 1 day [24 hours], 2 days [48 hours], etc.). We found that the largest chunk (41%) of adopters' purchases post app introduction occur within 48 hours of app use.

6.1. Temporal app use -- purchase linkages: The differences between app-linked and non-app-linked shoppers reveal that app-linked shoppers use the app for a greater number of days (Appendix Figure A3), longer periods in time (Appendix Figure A4), higher number of app sessions (Appendix Figure A5), and longer dwell time (Appendix

[8] In this analysis, we examine the six-month period after app launch and consider adopters in this period (54% of all adopters in the main sample) to analyze the changes close to the app launch date for explaining the short-term differences between app-linked and non-app-linked shoppers. Subsequently, we considered different time periods and found similar results.

Figure A6) during each session ($p < 0.001$) than non-app-linked shoppers in the six months after app launch.

Furthermore, app-linked shoppers use a greater number of app features across sessions. However, within a session, the number of features app-linked shoppers use does not significantly differ from those used by non-app-linked shoppers (Appendix Figure A7).[9] Extended use of features over a longer period of time might possibly explain subsequent purchases by app-linked shoppers better than the diversity of feature use within a session does. An alternative explanation can be that the number of sessions is correlated with purchases – those that shop more also use the app more times.

6.2. Geographical app use -- purchase linkages: Of the 30,026 app adopters who adopted the app in the first six months, 9,940 enabled location tracking in the app. Their median (mean) distance from the store of purchase when they access the app is 3.78 (22.39) miles. Moreover, 20.05% of their transactions take place after app use within a mile of the store of purchase, with 65.36% of the transactions occurring after app use within five miles. App-linked shoppers (average of 21.62 miles) tend to use the app much closer to stores ($p < 0.001$) than do non-app-linked shoppers (average of 30.96 miles).

6.3. Feature use -- purchase linkages: A comparison of features used by app-linked and non-app-linked shoppers shows that (a) within a session, app-linked shoppers

[9] We compute the within-session number of features used by calculating the average number of features a shopper clicks in the app in a given session (shopper-session level), while we calculate the across-session number of features through the cumulative sum of the number of features a shopper clicks across sessions in the six months post introduction. Therefore, this number can exceed the total number of features available in the app.

exhibit a higher frequency of use of features relating to offer list, product availability check, and call store ($p < 0.001$) and (b) across sessions, app-linked shoppers display a higher frequency of use of features relating to product search, product details, and rewards landing page ($p < 0.001$).

In addition to the comparison between app-linked and non-app-linked shoppers, we examine app-linked shoppers' features usage when they last accessed the app right before making a purchase as these features may be more directly related to their purchase decisions (see Figure A8 for the frequencies of shoppers using each feature before purchase). Registering for notifications, checking rewards landing page, and product details or search are the most commonly used features by shoppers before making purchases.[10]

Furthermore, for the subset of shoppers who share their locations, we break down the last-used features prior to an offline purchase based on (a) how far shoppers were from the store, and (b) how long they waited before they made their purchase, using quantile-splits. From Table A22, we find that (a) when shoppers are in close geographical and temporal proximity to the store, the rewards landing page feature is most used; (b) when shoppers are far from the store but are moderately far from their time of purchase (under 52.5 hours), they use the app as a product-information tool; (c) when shoppers are far in

[10] The rewards landing page displays the shopper's loyalty points. Clicking further on the rewards button on this page leads shoppers to browse rewards and use them in exchange for coupons (e.g., $5 off a purchase), redeemable in the store. Push notifications relate to special offers that are archived in a "message center" within the app.

both geographical and temporal dimensions, registering for notifications is a commonly used feature.

Two potential links between the app and subsequent purchases based on app use patterns are: in-store redemption of rewards discoverable through the app and product information available in the app. To test the rewards-based explanation, we examine differences in rewards-based purchases for those who used rewards in the last use of the app prior to purchase with those who did not. We find that the frequency, quantity, and value of rewards-based purchases are higher for those who used rewards in the app prior to purchase than for those who did not ($p < 0.001$). At the same time, non-rewards purchases are also higher for those who used rewards in the app prior to purchase than those who did not ($p < 0.001$). The results of t-tests comparing reward-feature users with non-users appear in Table A23. Both rewards purchases and non-rewards purchases are higher in dollar value for those who use the rewards feature in the app.

To test the product-information based explanation, we examine the number of unique SKUs and categories bought. For offline purchases, if shoppers use the app close to or inside a store, we should see additional cross-selling effects through additional in-store product exposure. From Table A24, we find that shoppers diversify their purchase baskets by buying a greater number of unique SKUs post app introduction, in particular, less popular items.[11] App adopters buy 16.4 different products relative to non-adopters,

[11] The number of observations in this analysis is 1,983,596 because it is conditional on making a purchase, accounting only for shopper-month observations when there is a purchase. Such a sample offers high interpretability as the number of SKUs, categories, or popularity of products purchased matter only when there is a purchase.

who purchase only 8.98 unique products ($p < 0.001$). Thus, product information appears to be a key benefit of the app.

6.4. App use – returns linkages: What could explain the higher frequency, quantity, and monetary value of product returns for app adopters relative to non-adopters? Were certain types of products, prompted by the use of the app, more likely to be returned? To answer these questions, we first examine the app features used before the original purchase that was subsequently returned. Shoppers who last used the offers feature in the app returned those purchases on 21% of the occasions. Presumably, purchasing a product at the lure of an offer in the app may be associated with subsequent product returns. On the other hand, shoppers who last used the product feature in the app, returned the product on only 12% of the occasions. Thus, purchases followed by the use of product-related features are less prone to returns possibly due to information effect present while using product-related features relative to the offer feature.

Next, we examine changes in returns based on the channel of purchase. We find that the frequency, quantity, and monetary value of returns of purchases in stores is greater than those for online purchases (Appendix Table A25). Counter-intuitively, store purchases are more likely to be returned by app adopters than non-adopters.

Interestingly, shoppers' trips to the store to make returns are often accompanied by new purchases particularly for app users. Overall, about 36.62% of the transactions involving product returns exhibited a purchase. The average monetary value of returns in these "returns transaction" is $57.79, while the average monetary value of purchases is $66.78. The net benefit (monetary value of new purchase minus returns during a "returns trip") is higher for the treated group than for the control group. Specifically, in the post app introduction period, the difference in the monetary value of new purchases made on the return trip and the monetary value of returns is \$3.81 for the app adopters compared with \$2.97 for the non-adopters ($p < 0.001$), suggesting higher incremental gains from returning adopters.

Managerial Implications

Our results offer key managerial implications. First, based on the difference in net monetary value of purchases for app adopters relative to non-adopters from the propensity score matched regression model with selection correction (\$21.11) and the most conservative estimate from the robustness checks (\$10.21) across the retailer's 55,580 adopters, we estimate that the retailer's net annual revenue increase after app introduction is approximately \$(6.8-14.1) million from these adopters.

Second, the finding that app adopters' purchases expand in both the online and offline channels suggests that managers should plan for app adopters visiting both the physical and online stores more often. Managers should expect more purchases from shoppers based on the time, location, and features used in the app of the shoppers. As a result, integrating individual-level shopper insights from the app into in-store and online experiences may help create additional value in these channels. Because shoppers are likely to access product information through the app, store associates can go beyond offering the standard product information.

Third, the finding that app adopters have greater product returns requires managers to proactively monitor returns from app adopters and devise interventions to keep product returns in check. Managers should also leverage app adopters' return visits and create opportunities for selling more items. They could remind returning shoppers of new product launches, promoted items, and rewards points or offers. In this way, managers could enhance the likelihood of returning shoppers purchasing new items on their return transactions.

Finally, managers should expect higher returns for products originally purchased in stores and rewards-based purchases than for products bought online and non-rewards purchases. This counterintuitive finding suggests that managers could use incentives such as free or reduced shipping to promote online purchases without worrying about potentially high returns from those online purchases.

Conclusion, Limitations, and Extension

In this chapter, we addressed three research questions on the relationship between a retailer's app introduction with multichannel shopping. We used non-experimental variation in the data from a large-scale retailer, combined with econometric techniques for mitigating self-selection.

Our analyses offer several key insights that demonstrate novel results and extend extant research. First, mobile app adopters have a higher frequency, quantity, and monetary value of both purchases and returns relative to non-adopters. App adopters buy 33% more frequently, 34% more quantity, and spend 37% more compared with non-adopters in the period after app introduction. At the same time, they return 35% more frequently, 35% more quantity and 41% more in dollar value. Overall, app adoption is associated with a 36% increase in net monetary value of purchases across all channels.

Second, both the online and offline purchases are higher for app adopters than for non-adopters. Third, the time, location, and features of app use provide descriptive evidence of how the app aids shopping in different channels. Specifically, app-linked shoppers (those who make a purchase within 48 hours of app use) are those who use the app when they are close to the store of purchase, and access the app for rewards, product details, and notifications. App adopters also purchase a more diverse set of items (including less popular products) than non-adopters. App adopters return products originally purchased in stores and through loyalty rewards more than those purchased online and without rewards.

Although our research is the first to quantify and explain the differences in app adopters and non-adopters of a retailer's mobile app with respect to a broad range of shopping outcomes, including returns, it has some limitations. First, our research relies on non-experimental variation in the data. Randomized field experiments, if feasible, could provide future researchers with unique opportunities for testing specific app-related manipulations. Second, our results capture the effects on the app adopters who use the app at least once, not those who downloaded but never used the app. Third, the number of cell towers serves as the source of exogenous variation in app adoption because it has little correlation with demand factors in our data and period. Despite our careful analysis, if there are any other unobservable confounds (especially if correlated with the number of cell towers), we need to be cautious about a causal interpretation of our estimates. The exogeneity of cell tower assumption will fall down in the presence of confounds correlated with both cell towers and demand factors, for example, in locations where the number of cell towers and population are highly correlated. Furthermore, our estimates are more representative of a local average treatment effect (LATE) than an average treatment effect (ATE) for the population. They are particularly relevant for app adopters whose decision to adopt the app is affected by the number of cell towers in the zip code (Sudhir and Talukdar 2015). Fourth, while we have examined the net monetary value of purchases, our data do not contain cost information. If cost data are available, it would be interesting to study the effect of app introduction on customer lifetime value. Fifth, we have data from only one retailer across channels. Future studies on mobile apps can examine data on multiple retailers to map shoppers' brand loyalty and preference resulting from the app. Likewise, future studies can examine these research questions in the context of other retailer types such as pure play retailers with a growing bricks-and-mortar presence (e.g., Warby Parker, Bonobos) and for mobile apps that allow in-app purchases. Finally, future research can also examine the marketing mix strategies for improving adoption of and engagement through apps.

CHAPTER IV

THE IMPACT OF MOBILE APP FAILURES ON PURCHASES IN ONLINE AND OFFLINE CHANNELS

Mobile commerce have seen tremendous growth in the last few years, with the majority of shopping journeys starting from a mobile device. About 2.4 billion people use smartphones worldwide. Mobile applications (henceforth, apps) have emerged as an important channel for retailers because they increase engagement and purchases across channels (e.g., Wang et al. 2015, Xu et al. 2016, Narang and Shankar 2019).

However, mobile apps can have a dark side for retailers. Failures experienced on retailers' mobile apps have the potential to negatively affect shoppers' engagement with the app and their shopping outcomes within the mobile channel. In addition, app failures may potentially have spillover effects across the different channels due to both substitution of purchases across channels and impact on preference for the retailer brand.

Preventing app failures is critical for managers because more than 60% shoppers abandon an app after experiencing failure(s) (Dimensional Research 2015). In 2016, app crashes were the leading cause of system failures, contributing 65% to all iOS failures (Blancco 2016). About 2.6% of all app sessions result in a crash, suggesting about 1.5 billion app failures across 60 billion app sessions annually (Computerworld 2014). Given these app failures and the potential damage these failures could create for firms' relationships with customers, determining the impact of app failures is important for formulating preventive and recovery strategies.

Despite the importance of app failures for firms and shoppers, not much is known about the impact of app failures on purchases. While app crashes in a shopper's mobile device have been shown to negatively influence subsequent engagement with the app (e.g., restart time, browsing duration, and activity level, Shi et al. 2017), the relationship between systemwide app failures and subsequent purchases has not been studied. Furthermore, a large proportion of shoppers use both online and offline retail channels. In such an omnichannel environment, it is critical for retailers to understand the impact of app failures on shopping behaviors not just within that channel, but also across channels (spillover effects). However, we do not yet know whether and how app failures influence purchases and spill over to other channels.

From a theoretical standpoint, the potential mechanisms behind such effects within and across channels are important. How much of these effects arise due to channel substitution post failure? What portion of the effects can be attributed to dilution of preference for the retailer's brand? Prior research has not addressed these interesting questions.

The effects of app failure may also differ across shoppers. Shoppers may be more or less negatively impacted by failures depending on factors such as shoppers' relationship with the firm (Goodman et al. 1995, Hess et al. 2003, Chandrashekaran et al. 2007, Knox and van Oest 2014, Ma et al. 2015) and shoppers' prior use of the firm's digital channels (Cleeren et al. 2013, Liu and Shankar 2015, Shi et al. 2017). It is important for managers to better understand how the effects of failure vary across shoppers so that they can devise better preventive and recovery strategies targeted at the shoppers. Yet, not much is known about heterogeneity in the effects of app failure on purchases across channels.

Our study fills these crucial gaps. We quantify and explain the impact of app failures on managerially important outcomes, such as the frequency, quantity and monetary value of purchases in the online and offline channels of a retailer. We address four research questions:

- What are the effects of a service failure in a retailer's mobile app on the frequency, quantity, and monetary value of subsequent purchases by the shoppers?
- What are the effects of a service failure in an app on purchases in the online and offline channels?
- What potential mechanisms explain the effects of a service failure in an app on purchases?
- How do these effects vary across shoppers or what factors moderate these effects?

Estimation of the effects of app failures on shopping outcomes is challenging. It is typically hard to estimate the impact of app failures on shopping behavior using observational data due to the potential endogeneity of app failures. This endogeneity may stem from an activity bias in that shoppers who use the app more frequently are more likely to experience failures than shoppers who use the app less frequently. Therefore, failure-experiencing shoppers may differ systematically from non-failure experiencers in their shopping behavior, leading to potentially spurious correlations between failures and shopping behavior. Panel data may not necessarily mitigate this issue because time-varying app usage/shopping activity is potentially correlated with time-varying app failures for the same reason. That is, shoppers are likely to engage more with the app when they are likely to purchase, potentially leading to more failures than in periods when shoppers engage less with the app. Additionally, the nature of activity on the app may be correlated with failures. For instance, a negative correlation between failures and purchases may result from a greater incidence of failures on the app's purchase page than on other pages. Thus, it is hard to make the case that correlations between app failures and shopping outcomes in observational data have a causal interpretation.

The gold standard among the methods available to uncover the causal impact of service failures is a randomized field experiment. However, such an experiment would be impractical in this context because a retailer will unlikely deliberately induce failures in an app even for a small subset of its shoppers for ethical reasons. Alternatively, we can use an instrumental variable approach to control for endogeneity. However, it is hard to come up with valid instrumental variables that exhibit sufficient variation to address the endogeneity concerns in this context.

We overcome the estimation challenges and mitigate the potential endogeneity of app failures using the novel features of a unique dataset from a large omnichannel retailer of video games, consumer electronics and wireless services. We exploit incidences of systemwide exogenous failure shocks in the retailer's mobile app due to server errors to estimate the causal effects of app failures. App users who logged into the app on the day of the failure were randomly exposed to the failures depending on whether they logged in during the time window of the exogenous shock. We estimate the effects of app failures using a difference-in-differences procedure that compares the pre- and post- failure outcomes for the failure experiencers with those of failure non-experiencers. Through a series of robustness checks, we confirm that failure non-experiencers act as a valid control for failure experiencers, providing us the exogenous variation to find causal answers to our research questions.

We investigate the potential mechanisms and moderators of the effects of failures on shopping behavior by exploiting the panel nature of our dataset. We test for the moderating effects of factors such as relationship with the firm and prior digital channel use on the effects of service failures. These factors have been explored for services in general (e.g., Ma et al. 2015, Hansen et al. 2018) but not in the digital or mobile context. In addition, we recover the heterogeneity of effects at the individual level using data-driven machine learning methods.

Our results show that app failures have a significant overall negative effect on shoppers' frequency, quantity, and monetary value of purchases across channels, but the effects are heterogeneous across channels and shoppers. A significant decrease in app engagement (e.g., number of app sessions, dwell time, and number of app features used) post failure explains the overall drop in purchases. Interestingly, the overall decreases in purchases across channels are driven by purchase reductions in brick-and-mortar stores, rather than in digital channels. Brand preference dilution after app failure explains the fall in store purchases, while channel substitution post failure explains the preservation of purchases in digital channels. Shoppers experiencing the failure when they are farther away from purchase (e.g., browsing information) experience greater negative effects of a failure than those closer to purchase (e.g., checking out). Surprisingly, the basket size and value of purchases rise for a small group of shoppers who were close to a store at the time of app failure. Furthermore, shoppers with less recent purchases, who are older and female, and those with a greater number of stores in their neighborhood are less sensitive to app failures. Finally, most shoppers (92%) react negatively to failures, but about 51% of these shoppers contribute to about 70% of the losses in annual revenues that amount to $(12-16.8) million or (35-48)% of the net income for the retailer in the data.

In the remainder of the chapter, we first discuss the literature related to service failures, cross-channel spillovers and consumer interaction with mobile apps, and develop the conceptual framework for our empirical analysis. Next, we discuss the data in detail, summarizing them and highlighting their unique features. Subsequently, we describe our empirical strategy, layout and test the key identification strategy, and conduct our empirical analysis of the effects of app failures. We then conduct robustness checks to rule out alternative explanations. We conclude by discussing the implications of our results for managers.

Conceptual Background and Related Literature

Services Marketing and Service Failures

The nature of services has evolved considerably since academics first started to study services marketing. For long, the production and consumption of services remained inseparable primarily because services were performed by humans. However, of late, technology-enabled services have risen in importance, leading to two important shifts (Dotzel et al. 2013). First, services that can be delivered without human or interpersonal interaction have grown tremendously. Online and mobile retailing no longer require shoppers to interact with human associates to make purchases. Second, closely related to this idea is the fact that services are increasingly powered by technologies such as mobile apps that allow anytime-anywhere access and convenience.

With growing reliance on technologies for service delivery, service failures are becoming more common. A service failure can be defined as service performance that falls below customer expectations (Hoffman and Bateson 1997). Service failures are widespread and are expensive to mend. Service failures resulting from deviations between expected and actual performance damage customer satisfaction and brand preference (Smith and Bolton 1998). Post-failure satisfaction tends to be lower even after a successful recovery and is further negatively impacted by the severity of the initial failure (Andreassen 1999, McCollough et al. 2000). In interpersonal service encounters, the human element and employee behaviors influence both failure effect and recovery (Bitner et al. 1990, Meuter et al. 2000). In technology-based encounters, such as those in e-tailing and with self-service technologies (e.g., automated teller machines [ATMs]), the opportunity for human interaction is typically none after experiencing failure (Forbes et al. 2005; Forbes 2008). However, there may be significant heterogeneity in how consumers react to service failures (Halbheer et al. 2018).

In the mobile context, specifically for mobile apps, it is difficult to predict the direction and extent of the impact of a service failure on shopping outcomes. First, mobile apps are accessible at any time and in any location through an individual's mobile device. On the one hand, because a shopper can tap, interact, engage, or transact multiple times at little additional cost on a mobile app, the shopper may treat any one service failure as acceptable without significantly altering her shopping outcomes. Such an experience differs from that with a self-service technological device such as an ATM, which may need the shopper to travel to a specific location. On the other hand, because a typical shopper uses multiple apps and can easily compare her experiences across them, a service failure in any one app may aggravate the shopper's frustration with the app, leading to strong negative effects on outcomes such as purchases from the relevant app provider.

Second, a mobile app is one of the many touchpoints available to shoppers in today's omnichannel shopping environment. Thus, a shopper who experiences a failure in the app could move to the web-based channel or even the offline or store channel. In such cases, the impact of a failure on the app could be zero or even positive (if the switch to the other channel leads to greater engagement of the shopper with the retailer). By contrast, if the channels act as complements or if the failure impacts the preference for retailer brand, a failure in one channel could impede the shopper's engagement in other channels. Thus, it is difficult to predict the effects of app failure, in particular, about how they might spill over to other channels.

Channel Choice and Channel Migration

A shopper's experience in one channel can influence her behavior in other channels. Prior research on cross-channel effects is mixed, showing both substitution and complementarity effects, leading to positive and negative synergies between channels (e.g., Avery et al. 2012; Pauwels and Neslin 2015). The relative benefits of channels determine whether shoppers continue using existing channels or switch to a new channel (Ansari et al. 2008; Chintagunta et al. 2012). When a bricks-and-clicks retailer opens an offline store or an online-first retailer opens an offline showroom, its offline presence drives sales in online stores (Wang and Goldfarb 2017, Bell et al. 2018). This is particularly true for shoppers in areas with low brand presence prior to store opening and for shoppers with an acute need for the product. However, the local shoppers may switch from purchasing online to offline after an offline store opens, even becoming less sensitive to online discounts (Forman et al. 2009). In the long run, the store channel shares a complementary relationship with the Internet and catalog channels (Avery et al. 2012).

While the relative benefits of one channel may lead shoppers to buy more in other channels, the costs associated with one channel may also have implications for purchases beyond that channel. In a truly integrated omnichannel retailing environment, the distinctions between physical and online channels blur, with the online channel representing a showroom without walls (Brynjolfsson et al. 2013). Mobile technologies are at the forefront of these shifts. More than 80% shoppers use a mobile device while shopping even inside a store (Google M/A/R/C Study 2013). As a result, if there are substantial costs associated with using a mobile channel (e.g., app failures), such costs may spill over to other channels. However, if shoppers treat the channels as substitutes, failures in one channel may drive the shoppers to purchase in another channel. If an app failure dilutes shoppers' preference for the retailer brand, it may lead to negative consequences across channels. Overall, the direction of the effect of app failures on outcomes in other channels such as in brick-and-mortar stores and online channels depends on which of these competing and potentially co-existing mechanisms is dominant.

Mobile Apps

The nascent but evolving research in mobile apps shows positive effects of mobile app channel introduction and use on engagement and purchases in other channels (Kim et al. 2015, Xu et al. 2016, Narang and Shankar 2019) and for coupon redemptions (Andrews et al. 2015, Fong et al. 2015, Ghose et al. 2018) under different contingencies.

To our knowledge, only one study has examined crashes in a mobile app on shoppers' app use. Shi et al. (2017) find that while crashes have a negative impact on future engagement with the app, this effect is lower for those with greater prior usage experience and for less persistent crashes. However, while they look at subsequent engagement of the shoppers with the mobile app, they do not examine purchases. Thus, our research adds to Shi et al. (2017) in several ways. First, we exploit the random variation in systemwide failures to determine the causal effects of failure. Second, we quantify the value of app failure's effects on subsequent purchases. Our outcomes include the frequency, quantity, and value of purchases, while the key outcome in that study is app engagement. Third, we examine the cross-channel effects of mobile app failures, including in physical stores, while Shi et al. (2017) study subsequent engagement with the app provider only within the app. Finally, we explore the mechanisms behind the effects of failure , and examine the moderating effects of relationship with the retailer and prior digital and heterogeneity in shoppers' sensitivity to failures using a machine learning approach.

Unlike prior related studies, our study (1) focuses on the effect of app failure on purchases, (2) quantifies the effects on multiple outcomes such frequency, quantity, and monetary value of purchases, (3) addresses the outcomes in each channel and across all channels (substitution and complementary effects), and (4) uncovers the mechanisms behind and moderators of the effects of app failure on shopping outcomes and heterogeneity in effects across shoppers.

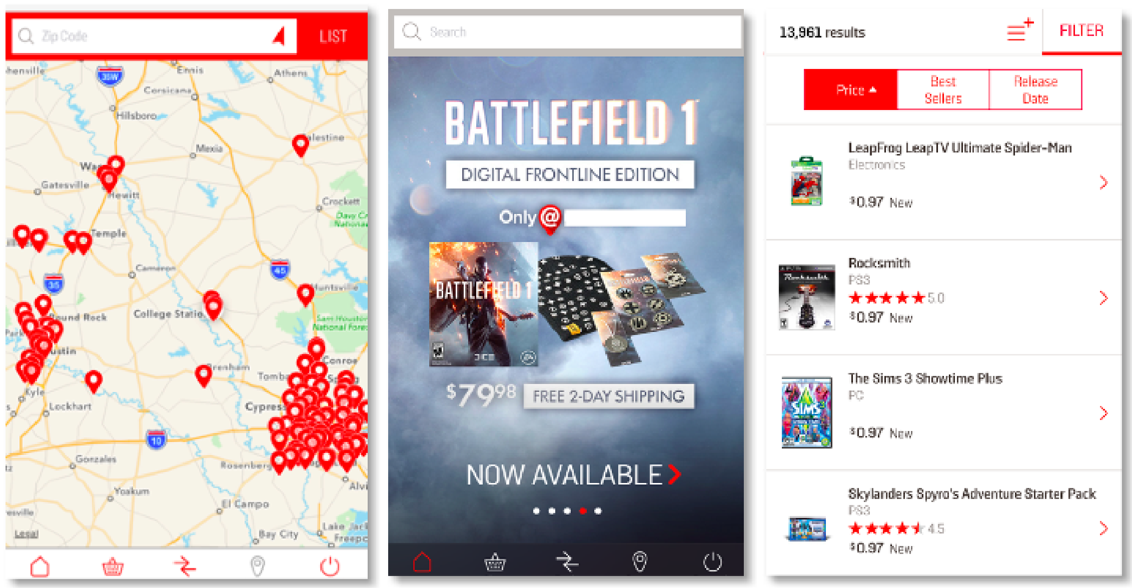
Research Setting and Data

We obtained the dataset for our empirical analysis from a large U.S.-based retailer. In the following paragraphs, we describe the retailer, the mobile app, and the channel sales mix.

The retailer sells a variety of products, including software such as video games and hardware such as video game consoles and controllers, downloadable content, consumer electronics and wireless services with 32 million customers. The gaming industry is large ($99.6 billion in annual revenues), and the retailer is a major player in this industry, offering us a rich setting. The retailer is similar to Walmart, PetSmart, or any other brick-and-mortar chain with a relatively large offline presence. The retailer's primary channel is its store network comprising 4,175 brick-and-mortar stores across the U.S. Additionally, it has a large ecommerce website.

The app allows shoppers to browse the retailer's product catalog, get deals, order online through mobile browser, locate nearby stores, as well as buy through the app. The app is typical of mobile apps of large retailers (e.g., PetSmart, Costco). The growth in adoption of the app is also similar to that of many large retailers. App adoption rate starts small and grows over time. Within four years of launch of the app, the number of app logins increased by 67%, reflecting the increased popularity of the app. Furthermore, about 26% of the shoppers bought online in the 12 months before the failure event we study. Figure 4 shows some screenshots from the app.

Figure 4. App Screenshots



The online and offline channel sales mix of the retailer in our data is typical of most large retailers. About 76% of the total sales for the top 100 largest retailers in the U.S. originated from similar retailers with a store network of 1,000 or more stores (National Retail Federation 2018). Most large retailers have a predominantly high brick-and-mortar presence. For these retailers, most of the transactions and revenues come from the offline channels, while online sales exhibit rapid growth. For example, Walmart's online revenues constitute 3.8% of all revenues, Costco's online sales are 3.6% of all sales, 1.3% of all PetSmart's sales come from the online channel, Home Depot generates 6.8% of all revenues from ecommerce, and 5.4% of Target's sales are through the online channel.[12] For the retailer in our data, online sales comprised 10.2% of overall revenues, greater than that for similar large retailers. Its online sales displayed a 13% annual average growth in the last five years, similar to these retailers who also exhibited double digit growth (Barron's 2018). Its online sales revenues are also substantial at $1.1 billion. We can see substantial cross-channel effects of a mobile app failure in the data. Thus, the impact of an app failure both the online and offline channels is important to consider.

Data and Sample

We study the impact of a systemwide failure that occurred on April 11, 2018.[13] The firm provided us with mobile app use data and transactional data across all channels for all the app users who logged into the app on the failure day. The online channel represents purchases at the retailer's website, including those using the mobile browser. Nested within the app use data are data on events that shoppers experience with their timestamps. The mobile dataset recorded the app failure event as 'server error.' Thus, this event represents an exogenous app breakdown, and the data allow us to identify shoppers who logged in to experience the systemwide app failure.

---

[12] Source: eMarketer Retail, https://retail-index.emarketer.com/

[13] We verified that these failures were systemwide and exogenous through our conversations with company executives. The failure on April 11, 2018 is farther away from the holiday season and helps rule out concerns about any potential idiosyncratic effects around Black Friday or other holiday season promotions.

To ensure that our estimates are not idiosyncratic to the single failure event, we additionally collected data on three other similar failures from 2014. Each failure is random and is totally exogenous. None of the failures coincided with marketing or promotion activities. The failures lasted for an average of five to 11 hours (11 for the October 2014 failure, five for the November 2014 failure, six for the December 2014 failure, and two for the April 2018 failure). They took place on different days of the week, including Monday, Wednesday, and Friday. The retailer experienced about 5-7 systemwide failures annually between 2014 and 2018, with most of them lasting longer than two hours. However, the data on the other failures are unavailable.

Table 14 provides the descriptive statistics for the variables of interest. Over a period of 14 days pre- and 14 days post- failure, shoppers make an average of approximately one purchase comprising two items valued at $43. In the 12 months preceding the failure, shoppers make an average of 11 purchases worth $598. Overall, 52% of the shoppers experience the failure.

Table 14. Summary Statistics

Variable	Mean	St. Dev.	Min.	Max.
Frequency of purchases	0.82	1.34	0	24
Quantity of purchases	1.61	3.32	0	270
Value of purchases ($)	43.31	96.42	0	2,144.89
App Failure (F)	0.52	0.50	0	1
Past frequency of purchases	11.9	13.43	0	493.0
Past quantity of purchases	24.78	33.48	0	1,600
Past value of purchases ($)	597.5	695.98	0	61,137.8
Past online purchase or not	0.25	0.43	0	1

Notes: These statistics are averaged over the pre and post- 14 days of the failure. N = 273,378.

Empirical Strategy

Overall Empirical Strategy

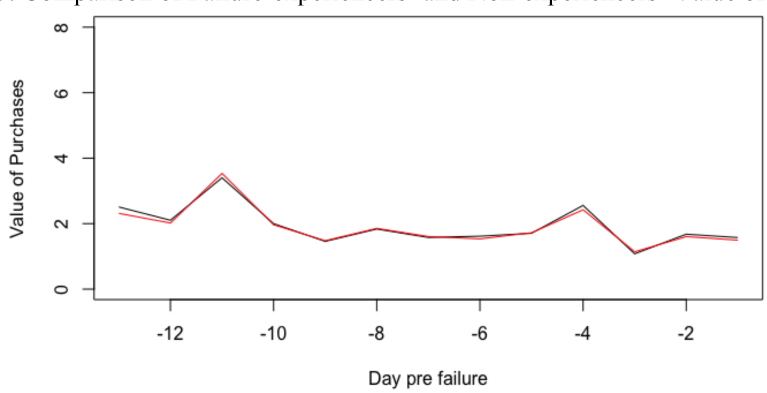
As outlined earlier, we leverage the natural experiment data by exploiting the exogenous systemwide failure. The main idea behind our empirical approach is that the usage of the app on a given day is random, and conditional on the usage of the app on the day of the failure, the experience of a failure by a specific shopper is random. We test this assumption in the data using a set of pre-failure variables. We find no systematic difference between shoppers who experienced failures and those who did not. We conduct a difference-in-differences analysis, comparing the post-failure with the pre-failure behaviors of shoppers who logged in on the day of the failure and experienced it (akin to a treatment group) relative to those who logged in on that day but did not experience the failure (akin to a control group).

To analyze the treatment effects within and across channels, we repeat this analysis with the same outcome variables separately for the offline and online channel. To understand the underlying mechanisms for the effects, we examine two explanations, brand preference dilution and channel substitution, using the data on shoppers' app engagement, closeness to purchase, location at the time of failure, and time to next purchase. To analyze heterogeneity in treatment effects, we first perform a moderator analysis using a priori factors such as prior relationship strength and digital channel use, followed by a data driven machine learning (causal forest) approach to fully explore all sources of heterogeneity across shoppers. Finally, we carry out multiple robustness checks.

Exogeneity of Failure Shocks

To verify the exogeneity of the failure shocks, we examine two types of evidence. First, we present plots of the behavioral trends in shopping for both failure-experiencers and non-experiencers for the failure shock in the 14 days before the app failure. Figure 5 depicts the monetary value of daily purchases by those who experienced the failure and those who did not. The purchase trends in the pre-period are parallel for the two groups ($p > 0.10$). The patterns for the frequency and quantity of purchases, and past online purchases are similar.

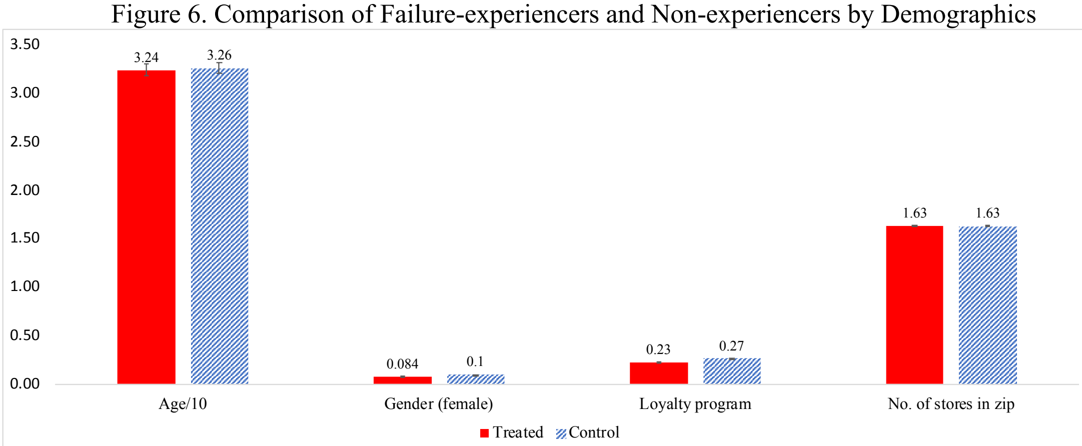
Figure 5. Comparison of Failure-experiencers' and Non-experiencers' Value of Purchases



Note: The red line represents the treated (failure experiencers) group, while the solid black line represents the control (failure non-experiencers) group.

Second, we compare the failure experiencers with non-experiencers across observed demographic variables, such as age, gender, membership in loyalty program, and number of stores in the shopper's neighborhood (see Figure 6). We do not find any significant differences in these variables across the failure experiencers and non-experiencers (p > 0.10).

Figure 6. Comparison of Failure-experiencers and Non-experiencers by Demographics



Notes: Gender = 0 (male), 1 (female); loyalty program level represents whether shoppers were enrolled (=1) or not (=0) in an advanced reward program with the retailer on the day of failure.

Econometric Model and Identification

To estimate the effects of app failure on shopping outcomes, we rely on a quasi-experimental research design with a difference-in-differences approach (e.g., Angrist and Pischke 2009) that leverages a systemwide failure shock and compares app users who experience this shock with those who do not, given that they accessed the app on the day of the failure. Note that we have already shown that the two groups do not differ systematically either on demographics or on behavioral variables in the period leading up to the app failure, offering face-validity.

Our two-period linear difference-in-differences regression takes the following form:

(5) $Y_{it} = \alpha_0 + \alpha_1 F_i + \alpha_2 P_t + \alpha_3 F_i P_t + \vartheta_{its}$

where i is the individual, t is the time period, and Y is the outcome variable, F is a dummy variable denoting treatment (1 if shopper i is experienced the systemwide app failure s and 0 otherwise), P is a dummy variable denoting the period (1 for the period after the systemwide app failure s and 0 otherwise), α is a coefficient vector, and ϑ is an error term. We cluster standard errors at the shopper level (Bertrand et al. 2004). The coefficient of F_iP_t, i.e., α_3 is the treatment effect of the app failure.[14]

The assumptions underlying the identification of this treatment effect are: (1) the failure is random conditional on a shopper logging into the app during the time window of the failure shock, and (2) the change in outcomes for the non-failure experiencing app users is a valid counterfactual for the change in outcomes that would have been observed for failure-experiencing app users in the absence of the failure.

Empirical Analysis Results

Relationship between App Failures and Purchases: Model-free Evidence

We first examine the overall differences in post-failure behaviors between shoppers who experienced failures and those who did not using model-free evidence 14 days pre and post failure. We choose a 14-day window period for two main reasons. First, this window is bimonthly and close to the mean interpurchase time of 16 days in our dataset. Second, it is not too long to include the occurrence of other events, including other systemwide failure shocks.[15]

---

[14] Because we analyze the short-term effect of a service failure (14 and 30 days), we do not have multiple observations per shopper post failure for us to include shopper fixed effects in our analysis.
[15] We also estimated a model with dynamic treatment effects for a longer period of five weeks pre- and post- the failure shock and found similar and persistent effects (see Figure 4 and Table D1).

Table 15 reports the model-free results for both failure experiencers (70,568 treated) and non-experiencers (66,121 control) given that they accessed the app on the day of the failure. We find that for post failure, shoppers who experienced the systemwide failure had 0.04 ($p < 0.01$) lower purchase frequency, 0.07 ($p < 0.01$) lower purchase quantity, and $2.42 ($p < 0.01$) lower monetary value than shoppers who did not experience the failure. A simple comparison of shopping outcomes across the two groups shows that the average monetary value of purchases increased by 81.8% ($30.41 to $55.28) for failure-experiencers, while it increased by 87.6% ($30.75 to $57.70) for non-failure experiencers post failure relative to the pre period ($p < 0.01$).[16] Given our identification strategy, the diminished growth in the monetary value of purchases for failure experiencers relative to non-experiencers comes from the exogenous failure shock.

Table 15. Model-free Evidence: Means of Outcome Variables for Treated and Control Groups

Variable	Treated pre Period	Treated post period	Control pre period	Control post period
Frequency of purchases	0.74	0.89	0.75	0.93
Quantity of purchases	1.52	1.69	1.52	1.76
Value of purchases ($)	30.41	55.28	30.75	57.70
Frequency of purchases – Online	0.03	0.04	0.03	0.04
Quantity of purchases – Online	0.05	0.06	0.05	0.07
Value of purchases – Online ($)	1.34	2.93	1.50	3.17
Frequency of purchases – Stores	0.70	0.85	0.71	0.88
Quantity of purchases – Stores	1.47	1.63	1.47	1.69
Value of purchases – Stores ($)	29.07	52.35	29.25	54.53

Notes: These statistics are based on pre- and post- 14 days of the failures. N = 273,378.

[16] Increasing sales trend between the pre- and post- period for both the groups is partially due to the April 19 weekend in the post period that witnessed the release of a new game.

Main Diff-in-Diff Model Results

The results from the difference-in-differences model in Table 16 show a negative and significant effect of app failure on the frequency (α_3 = -0.024, p < 0.01), quantity (α_3 = -0.057, p < 0.01), and monetary value of purchases (α_3 = -2.181, p < 0.01) across channels. Relative to the pre-period for the control group, the treated group experiences a decline in frequency of 3.20% (p < 0.01), quantity of 3.74% (p < 0.01), and monetary value of 7.1% (p < 0.01).[17]

Table 16. DID Model Results of Failure Shocks for Purchases across Channels

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.024** (0.008)	-0.057** (0.02)	-2.181** (0.681)
Failure experiencers	-0.021** (0.007)	-0.03 (0.018)	-0.694* (0.302)
Post shock	0.178*** (0.006)	0.236*** (0.014)	26.947*** (0.497)
Intercept	0.75*** (0.005)	1.523*** (0.012)	30.755*** (0.219)
R squared	0.018	0.001	0.004
Effect size	-3.20%	-3.74%	-7.09%

Notes: Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, ** p < 0.01, * p < 0.05. N = 273,378. DID = Difference-in-Differences.

Next, we examine the channel spillover effects of app failures in greater depth. We split the total value of purchases into store-based purchases and online purchases. Table

[17] We calculate the percentage change by dividing the treatment coefficient of the variable by the intercept. For instance, the coefficient for frequency of purchase (0.024) divided by that in the pre-period (0.75) amounts to a 3.2% change.

17 reports the results for these alternative channel-based dependent variables. There is a negative and significant effect of app failure on the frequency (α_3 = -0.02, p < 0.001), quantity (α_3 = -0.05, p < 0.001), and monetary value of purchases (α_3 = -2.18, p < 0.001) in the offline channel. Interestingly, we do not find a significant (p > 0.10) effect of app failure on any of the purchase outcomes in the online channel. Because there is no corresponding increase in the online channel and because the overall purchases drop, we conclude that the decreases in overall purchases across channels are largely due to declines in in-store purchases.

Table 17. DID Model Results of Failure Shocks for Purchases by Channel

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of Purchases
Failure experiencers x Post shock (DID)	-0.022** (0.008)	-0.055** (0.019)	-2.088** (0.66)	-0.001 (0.002)	-0.003 (0.003)	-0.093 (0.154)
Failure experiencers	-0.018** (0.006)	-0.025 (0.017)	-0.527 (0.293)	-0.003* (0.001)	-0.005* (0.002)	-0.167** (0.064)
Post shock	0.17*** (0.005)	0.221*** (0.014)	25.275*** (0.482)	0.009*** (0.001)	0.015*** (0.002)	1.672*** (0.113)
Intercept	0.714*** (0.005)	1.47*** (0.012)	29.255*** (0.213)	0.036*** (0.001)	0.054*** (0.002)	1.5*** (0.048)
R squared	0.0038	0.0001	0.0169	0.0016	0.0002	0.0003
Effect size	-3.08%	-3.74%	-7.14%			

Notes: Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, ** p < 0.01, * p < 0.05. N = 273,378. DID = Difference-in-Differences.

Mechanism Behind the Effects of Failures on Shopping Outcomes

We now provide descriptive evidence for the potential mechanism behind the results. The overall negative effect of app failure on shopping outcomes across channels could be due to decreases in intermediate outcomes such as shopping engagement after failure. To explore this possibility, we examine the effect of app failure on app engagement variables such as the number of app sessions, the average dwell time per session, and the average number of app features used in each session. The results of the corresponding DID model appear in Table 18. The DID coefficient for each of the three variables is negative and significant ($p < 0.001$), suggesting that app failure diminishes engagement. In addition, the results for the same variables for purchase related pages and sessions also show negative and significant effects ($p < 0.001$) of app failure. Purchase-related pages in an app involve actions closer to purchase, such as adding a product to shopping cart, clicking checkout, or making payment. In contrast, non-purchase-related pages involve actions farther from purchase, such as browsing products or store related information.

Table 18. DID Model Results for App Engagement Variables

	Dependent Variables		
Independent Variable	No. of app sessions	Average dwell time per session	Average no. of app features
Failure experiencers x Post shock (DID)	-0.689*** (0.005)	-7.444*** (0.072)	-4.678*** (0.024)
Failure experiencers	0.651*** (0.005)	7.041*** (0.065)	4.508*** (0.021)
Post shock	-0.624*** (0.003)	-3.525*** (0.047)	-2.558*** (0.016)
Intercept	0.727*** (0.003)	4.654*** (0.04)	3.067*** (0.013)
R squared	0.4211	0.1833	0.4843

Notes: Robust standard errors clustered by shoppers are in parentheses; the dependent variables are measured 5 hours pre- and post- failure; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. N = 273,378. DID = Difference-in-Differences.

The differential effect of app failure across the channels could be explained by two countervailing forces: channel substitution and brand preference dilution. The channel substitution effect occurs when app failure experiencers move to the mobile web browser, the desktop website, or the physical store to complete their intended purchase. Brand preference dilution happens when app failure experiencers are annoyed or dissatisfied with the retailer, leading to reduced value of purchases overall and in the store. It is possible that channel substitution effect is predominant for online shoppers, while brand preference dilution is prevalent for shoppers who purchase primarily in stores.

To better understand the differential effects of app failure across channels, we first examine the effect of app failure across shoppers who were close to and far from purchase at the time of failure. Table 19 reports the DID model results when the app failure occurred on purchase related and non-purchase related pages. The effects of app failure is negative and significant ($p > 0.01$) on all the outcome variables for shoppers who experience failure on a non-purchase related page than for shoppers who experience failure on a purchase related page. Shoppers who already have a strong purchase intent and are on a purchase-related page right before the failure are not as negatively affected as those without a strong purchase intent or on a non-purchase related page.

Table 19. DID Model Results for Failures Occurring on Purchase and Non-purchase Related Pages

	(1) Failure on purchase related page			(2) Failure on non-purchase related page		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of Purchases
Failure experiencers x Post shock (DID)	0.000 (0.013)	-0.016 (0.038)	0.907 (1.195)	-0.053*** (0.009)	-0.108*** (0.022)	-4.627*** (0.763)
Failure experiencers	-0.019 (0.011)	-0.016 (0.034)	-0.246 (0.52)	-0.041*** (0.007)	-0.075*** (0.019)	-1.208*** (0.341)
Post shock	0.178*** (0.006)	0.236*** (0.014)	26.947*** (0.497)	0.178*** (0.006)	0.236*** (0.014)	26.947*** (0.497)
Intercept	0.75*** (0.005)	1.523*** (0.012)	30.755*** (0.219)	0.75*** (0.005)	1.523*** (0.012)	30.755*** (0.219)
R squared	0.004	0.001	0.019	0.004	0.001	0.018

Notes: Robust standard errors clustered by shoppers are in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. DID = Difference-in-Differences. N = 160,662 for failure on purchase related page. N= 217,418 for failure on non-purchase related page.

Failure-experiencers who were close to a purchase or had purchase intent, would have had to determine whether to complete the transaction, and if so, whether to do it online or offline. For shoppers who typically buy online, the cost of going to the retailer's website to complete a purchase interrupted by the app failure is smaller than that of going to the store to complete the purchase. Therefore, these shoppers will likely complete the transaction online and not exhibit any significant decrease in shopping outcomes in the online channel post failure. Thus, channel substitution effect likely explains the behavior of such shoppers post app failure. In contrast, shoppers who typically buy in the retailer's brick-and-mortar stores and who experience the systemwide app failure, will likely have a diminished perception of the retailer and will have fewer incentives to buy from the brick-and-mortar stores in the future. Thus, the brand preference dilution effect may prevail for these shoppers after app failure.

To further explore the channel substitution effect, we examine the time elapsed between the occurrence of the failure and subsequent purchase in the online channel. Failure experiencers' inter-purchase time online ($Mean_{treated}$ = 162.8 hours) is much shorter than non-experiencers' ($Mean_{control}$ = 180.7 hours) ($p = 0.003$). This result suggests that after an app failure, shoppers look to complete their intended purchases in the online channel.

Finally, to more deeply understand channel substitution effect, we also examine the effect of the failure for shoppers who were close to a physical store at the time of app failure. Shoppers who are closer to a store when they experience the app failure could more easily complete their purchase in the store than shoppers farther from a store. Table 20 reports the DID model for the subsample of shoppers located within two miles of the retailer's store at the time of failure. The results show that shoppers closer to the store are not negatively affected by the failure ($p < 0.05$). Rather surprisingly, both the basket size and the monetary value of purchases for shoppers close to a store are significantly higher after app failure ($p < 0.05$). This result suggests that shoppers who visit a physical store after an app failure to complete their purchase end up buying additional items in the store. Thus, app failure has an unintended positive effect on such shoppers. However, the proportion of such shoppers in the sample is very small (2.4%), so the overall effect of app failure is still negative.

Table 20. DID Model Results for Value of Purchases and Basket Size by Channel for Shoppers Close to a Store (< 2 miles) at the Time of Failure

	Offline		Online	
Variable	Value of purchases	Basket Size	Value of Purchases	Basket Size
Failure experiencers x Post shock (DID)	13.542* (5.307)	0.218* (0.104)	0.885 (1.178)	0.02 (0.039)
Failure experiencers	2.419 (2.171)	0.108 (0.085)	-1.096* (0.479)	-0.003 (0.026)
Post shock	37.833*** (3.15)	0.002 (0.006)	2.382** (0.882)	0.009*** (0.002)
Intercept	32.458*** (1.336)	0.791*** (0.005)	2.251*** (0.39)	0.042*** (0.001)
R squared	0.0395	0.0001	0.0027	0.0002

Notes: Robust standard errors clustered by shoppers are in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. N = 6,572. DID = Difference-in-Differences. Two miles is the median distance from store at the time of failure.

We also examine the heterogeneity in the sensitivity of shoppers to app failures in two ways. We use a theory-based moderator approach as well as a data-driven machine learning approach.

Moderators: Relationship Strength and Prior Digital Use

The literatures on relationship marketing and service recovery suggest two factors may moderate the impact of app failures on outcomes: relationship strength and prior digital channel use.

Relationship Strength

The service marketing literature offers mixed evidence on the moderating role of the strength of customer relationship with the firm in the effect of service failure on shopping outcomes. Some studies suggest that stronger relationship with firms may aggravate the effect of failures on product evaluation, satisfaction, and on purchases (Goodman et al. 1995, Chandrashekaran et al. 2007, Gijsenberg et al. 2015). Other studies show that stronger relationship attenuates the negative effect of service failures (Hess et al. 2003, Knox and van Oest 2014).

Consistent with the direct marketing literature (Schmittlein et al. 1987, Bolton 1998), we operationalize customer relationship using RFM (recency, frequency, and monetary value) dimensions. Because of high correlation between the interactions of frequency with (failure experiencers x post shock) and value of purchases with (failure experience x post shock) ($r = 0.90$, $p < 0.001$) and because value of purchases is more important from the retailer's perspective, we drop frequency of past purchases and estimate a model with interactions of each of recency and value of past purchases with the interaction between failure experiencers and the post shock indicator ($F_i P_t$) in the proposed model.

Prior Digital Channel/Online Use/Experience

A shopper's prior digital channel/online use or experience with the retailer may moderate the effects of service failure on shopping outcomes. On the one hand, more digitally experienced app users may be less susceptible to the negative impact of an app crash on subsequent engagement with the app than less digitally experienced app users (Shi et al. 2017) because they are conditioned to expect some level of technology failures. This reasoning is based on the product harm crises literature (Cleeren et al. 2013, Liu and Shankar 2015) and the broader expectation-confirmation theory (Oliver 1980). The more experience a customer has with a service, the less impact a single piece of new information (from failure) will have on service evaluation and usage (Tax et al. 1998, Cleeren et al. 2008). On the other hand, prior digital exposure and experience with the firm may heighten shopper expectations and make them less tolerant of service failures. To examine the empirical moderating effect of prior digital channel use, we operationalize it as the cumulative number of purchases that the shopper made from the retailer's website prior to experiencing a failure.

The results of the model with relationship strength and past digital channel use as moderators appear in Table 21. Consistent with our expectation, the monetary value of past purchases has positive interaction coefficients with the DID variable across all outcome variables when significant ($p < 0.001$). Thus, app failures affect high value shoppers less. Similarly, recency has positive coefficients ($p < 0.001$), suggesting that the more recent shoppers are more tolerant of failure. A failure shock also affects the frequency, quantity, and value of purchases ($p < 0.001$) of shoppers with greater digital channel or online purchase exposure or experience with the retailer.

Table 21. DID Model Results of Failure Shocks for Purchases across Channels: Moderating Effects of Relationship with Retailer and Past Online Purchase

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.15*** (0.011)	-0.302*** (0.028)	-12.106*** (0.823)
DID x Past value of purchases	0.00*** (0.000)	0.000*** (0.000)	0.018*** (0.001)
DID x Recency of purchases	0.001*** (0.000)	0.001*** (0)	0.017*** (0.004)
DID x Past online purchase frequency	-0.012*** (0.002)	-0.021*** (0.004)	-1.066*** (0.129)
Past value of purchases	0.001*** (0.000)	0.001*** (0.000)	0.024*** (0.000)
Recency	-0.002*** (0.000)	-0.004*** (0.000)	-0.087*** (0.002)
Past online purchase frequency	-0.013*** (0.001)	-0.032*** (0.002)	-0.881*** (0.057)
Failure experiencers	0.013 (0.007)	0.052** (0.017)	0.750 (0.503)
Post	0.178*** (0.007)	0.236*** (0.017)	26.947*** (0.510)
Intercept	0.534*** (0.006)	0.972*** (0.015)	21.144*** (0.428)
R squared	0.1199	0.0957	0.0741

Notes: DID = Difference-in-Differences. Robust standard errors are in parentheses; cohort fixed effects included; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. N = 273,378. The number of observations includes observations of shoppers with at least one purchase in the past for computing recency.

Heterogeneity in Shoppers' Sensitivity to App Failures (Treatment Effect)

In addition to theoretically driven moderator variables, we are also interested to further explore heterogeneity in treatment effects relating to managerially useful additional observed variables (e.g., age, gender, membership in loyalty program) not fully investigated by prior research. Estimation of individual level treatment effects and knowledge of their drivers are useful for managers for service failure prevention and recovery purposes. Unfortunately, including these variables as additional moderators will impose a huge burden on the DID analysis as the number of potential main and interaction effects will become unmanageably large.

Recent advances in causal inference using machine learning allow us to recover individual-level conditional average treatment effects (CATE) (Athey and Imbens 2016, Athey et al. 2017, Wager and Athey 2018). These methods use ideas from regression trees and random forests (Breiman 2001) to identify subpopulations of the data that differ in the magnitude of the treatment effect. These methods have been applied in marketing to customer churn and information disclosure (Ascarza 2018, Guo et al. 2018). In our context, we estimate a causal forest model, an ensemble of causal trees that averages the predictions of treatment effects produced by each tree for thousands of trees.[18] We follow the causal tree estimation with a regression of CATE on the moderators and observed demographic variables.

The estimates from causal forests using 1,000 trees appear in Appendix Table B1. The average CATE is -1.91 for the test data, close to the overall average treatment effect on the treated (ATET) of -2.18 obtained from the regression model. Furthermore, 92% of the shoppers have a negative value of CATE; their average is -3.42. The distribution of CATE across shoppers appears in Appendix Figure B1. The shopper quintiles based on the levels of CATE reflects this distribution in Appendix Figure B2, which reveals that Segments 1 and 2 of the most sensitive shoppers also exhibit higher variance than the rest.

Next, we regress the CATE estimate on the covariate space to identify the covariates that best explain treatment heterogeneity. The results from such an ordinary least squares

[18] In Appendix D, we provide an overview of causal trees and describe the algorithm for estimating a single causal tree followed by bagging a large number of causal trees into a forest.

(OLS) regression of CATE on covariates appear in Appendix Table B2. All the covariates except mobile operating system are significant ($p < 0.001$). Shoppers with higher past purchase value, loyalty program membership, with higher frequency of past online purchases, and access to a greater number of stores are less sensitive to app failures than others. However, older and female customers and those with more recent purchases are more sensitive to failures than others.

Robustness Checks and Ruling out Alternative Explanations

We perform several robustness checks and tests to rule out alternative explanations for the effect of app failure on purchases. Because our goal is to examine the consistency of results over time and across multiple customer base, we analyzed a data from a pooled sample containing the April 2018 failure sample and three other systemwide app failures that occurred on October 10, November 3, and December 10 in 2014 for which we could obtain data.

Alternative model specifications

Although the failure in our data is exogenous, to be sure, in addition to our proposed difference-in-differences model, we also estimate models with propensity score matching and Poisson count data models for the frequency and quantity variables. The results from these models replicate the findings from Tables 16 and 17 and appear in the Appendix Tables C1-C2 and D1-D2, respectively. The coefficients of the treatment effect from Table C1 and D1 represent changes in outcomes due to app failures, conditioned on covariates through propensity scores. These results are substantively similar to those in Tables 16 and 17. The insensitivity of the results to control variables suggests that the effect of unobservables relative to these observed covariates would have to be very large to significantly change our results (Altonji et al. 2005). Similarly, the results are robust to a Poisson specification, reported in Tables C2 and D2.

Alternative time periods

In addition to results from 14 days pre- and post- app service failure, we present results for 30 days pre- and post- models in Appendix Tables C3 and D3. These results are substantively similar. Our proposed model leverages a shorter 14-day period to avoid overlaps across pre- and post- periods between shocks that occur close to each other (e.g., a 30-day post period for the November failure shock would overlap with a 30-day pre period of the December failure shock).

Outliers

We re-estimate the models by removing outlier (extremely high) spenders (three standard deviations in monetary value of purchases in the pre-period) from our data. Appendix Tables C4 and D4 report these results. We find consistent and even stronger results.

Existing shoppers

Another possible explanation for app failures' effect can be that only new or dormant shoppers are sensitive to failures, perhaps due to low switching costs. Therefore, we remove those with no purchases in the last 12 months to see if their behavior is similar to that of the existing shoppers. Indeed, Appendix Tables C5 and D5 report substantively similar results after excluding the new or dormant shoppers.

Alternative measures of digital channel use moderators

In lieu of past online purchases dummy as a measure of prior digital channel use, we use a measure based on median splits in the number as well as the share of online purchases. Furthermore, we also use another measure based on prior app usage in the time between app launch and each server failure in the app. The results for alternative online purchase measures are almost the same as our proposed model results, except for online purchases. Similar results emerge from the app use measure in Appendix Tables C6 and D6.

Regression discontinuity analysis

To ensure that there are no unobservable differences between failure experiencers and non-experiencers based on the time of login, we carry out a 'regression discontinuity' (RD) style analysis in the one hour before the start time of the service failure. For the RD analysis, we consider only app users in the neighborhood of this time, using as control group those users who logged in one hour before and after the failure period and as treated the users who logged in during the failure period. The results are substantively similar to our main model results and are reported in Appendix Tables C7 and D7.

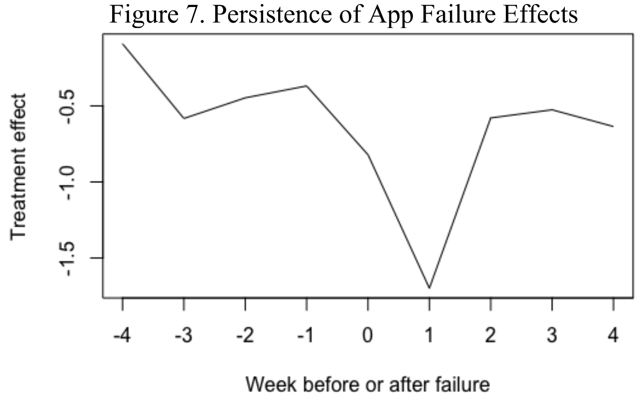
Multiple failures analysis

To ensure that the effects are robust to shoppers who experience more than one of the four failure shocks, we estimate the main model after including multiple failure experiencers in the sample. The results are similar to our main model results and appear in Appendix Tables C8 and D8 with the exception of frequency and quantity of online purchases that also decrease for those with multiple failures. The results also show that app failure decreases purchases in the online channel, suggesting that the presence of multiple failures might make shoppers more sensitive to failures in their online purchases as well.

Persistence of Failure

Our main analysis shows 14-day effects of app failures. To explore if these effects persist over longer periods of time, we examined the outcomes four weeks pre- and post-failure event. There is a steep fall in purchases in the period immediately after the failure. However, purchases climb back to higher levels over the next three weeks. Nevertheless, they return to levels lower than the pre-period average levels. As a result, the diminished impact of failure persists over time. These patterns appear in Figure 7 and Appendix Table E1. The table shows the coefficients for a DIFF-in-DIFF regression with weekly-aggregated data and interactions of weekly dummies with TREAT to map the coefficient path.

Figure 7. Persistence of App Failure Effects



Stacked model for channel effects

The results for online and offline purchases in Table 17 do not show the relative sizes of the effects across the two channels. To examine these relative effects, we estimate a stacked model of online and offline outcomes that includes a channel dummy. The results for this model appear in Appendix Table E2. We interpret the effects as a proportion of the purchases within the channel and conclude that the effects in the offline channel are more negative than those in the online channel ($p < 0.01$ or better).

Discussion

In this chapter, we addressed novel research questions: What is the effect of a service failure in a retailer's mobile app on the frequency, quantity, and monetary value of purchases in online and offline channels? What possible mechanisms may explain these effects? How do shoppers' relationship strength and prior digital channel use moderate these effects? How heterogeneous is shoppers' sensitivity to failures? By answering these questions, our research fills an important gap at the crossroads of three disparate streams of research in different stages of development; the mature stream of service failures, the growing stream of omnichannel marketing, and the nascent stream of mobile marketing. We leveraged a random systemwide failure in the app to measure the causal effect of app failure. To our knowledge, this is the first study to causally estimate the effects of digital service failure using real world data. Using unique data spanning online and offline retail channels, we examined the spillover effects of such failures across channels and examined heterogeneity in these effects based on channels and shoppers.

Our results reveal that app failures have a significant negative effect on shoppers' frequency, quantity, and monetary value of purchases across channels. These effects are heterogeneous across channels and shoppers. Interestingly, the overall decreases in purchases across channels are driven by reductions in store purchases and not in digital channels. Furthermore, we find that that shoppers with higher monetary value of past purchases are less sensitive to app failures.

Overall, our nuanced analyses of the mechanisms by which an app failure affects purchases offer new explanations that are insightful in a cross-channel context. In this way, our findings are consistent with the view that some customers may be tolerant and forgiving of technological failures (Meuter et al. 2000). Finally, our study offers novel insights into the cross-channel implications of app failures.

Managerial Implications

The effects of failures are sizeable for any retailer to alter its service failure preventive and recovery strategies. Based on our estimates, the economic impact of just one app failure in our data is a conservative estimate of annual revenue loss of about $2.4 million for the retailer.[19] The retailer experiences about 5-7 failures each year, resulting in an annual loss of $(12-16.8) million. This loss amounts to a substantial amount, about (35-48)% of the retailer's net income for 2017-2018.

[19] We compute this figure by using the weekly effect coefficients in Table D1, i.e., $(0.82 + 1.70 + 0.58 + 0.53)*N for the first four weeks and $(0.64)**48*N assuming that the fifth week's persistent effects remain for the future 48 weeks for N = 70,568 failure experiencers, totaling $2.4 million.

The economic effect is meaningful for several reasons. First, retailers operate on thin margins and are cost-conscious, so a loss of 35-48% of net income is impactful. Second, the effect size of 7.1% from our results is consistent with and even higher than those from other similar causal studies. For example, exposure to banner advertising has been shown to lift purchase intention by 0.473% worth 42 cents/click to the firm (Goldfarb and Tucker 2011). Goldfarb and Tucker (2011) argue, “although the coefficient may seem small, it suggests an economically important impact of online advertising.” Third, in the mobile context, the effect of being in a crowd (of five people relative to two per square meter when receiving a mobile promotion) results in 2.2% more clicks (Andrews et al. 2015). Fourth, as sales through the mobile app and online sales are growing rapidly, this effect is only getting larger. Finally, our estimates are for one two-hour app failure in a year. The retailer experienced about 5-7 systemwide failures annually between 2014 and 2018, with most lasting longer than two hours. The insights from our research better inform executives in managing their mobile app and channels and offer practitioner implications for service failure preventive and recovery strategies.

Preventive Strategy

The finding that app failures result in a 7.1% decrease in monetary value of purchases helps managers estimate the quantitative effect of an app failure. Managers can use this estimate to budget resources for their efforts to prevent or reduce app failures. The effects may be more pronounced for pure-play online retailers.

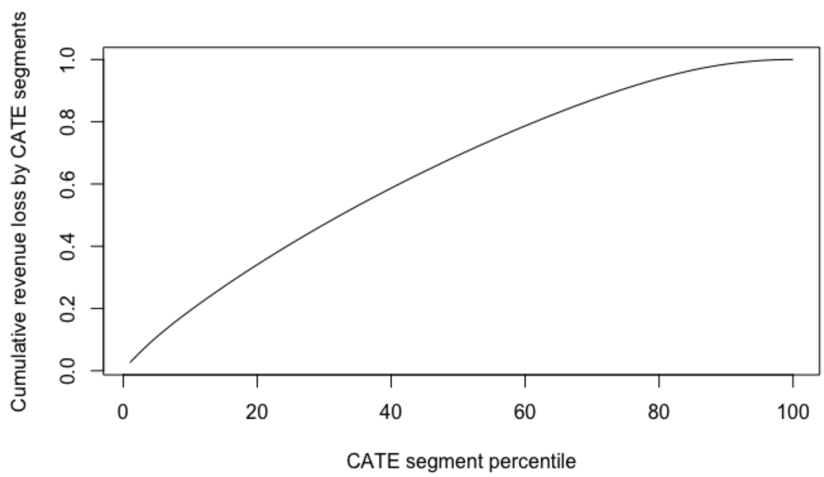
Managers should anticipate, monitor, regulate, and lower the likelihood of a systemwide failure that affects a majority of their customer base. Over time, these failures may be relatively easier to detect and timely remedial interventions through bug fixes and new app versions can minimize the effect of failure.

The result that the adverse effect of failure is lower for shoppers closer to purchase and purchasing online provides interesting pointers to retailers. In general, managers should encourage shoppers to use the app more and get closer to purchase and more habituated to purchasing through the app. They could offer one-time or limited-time incentives for those who have not clicked the checkout or purchase tabs in the app.

In particular, by identifying failure-sensitive shoppers based on relationship strength and prior digital use, managers can take proactive actions to prevent these shoppers from reducing their shopping intensity with the firm. They can issue customized assurances to these shoppers through other communication channels such as email or voicemail; warning them of likely disruptions in the app can preempt negative attributions and attitudes, and limit any impending damage to the brand and revenues due to app failure. Often random failures may be hard to predict or control, so managers should be ready when failures do occur. This is where knowing the failure-sensitive shoppers (based on the effects of past failures) ahead of time and reacting quickly after the failure is very useful for retailers.

Finally, using the individual-level CATE estimates, managers can target shoppers for preventive strategy. Figure 8 represents the loss of revenues (spending) from each percentile of shoppers at different levels of failure sensitivity. About 70% of the losses in revenues due to failure arise from just 51% of the shoppers. Managers can identify these most failure-sensitive shoppers and manage these shoppers' expectations through email and app notification messaging channels. They can use the results of moderation analysis and causal forest to identify potential new failure insensitive shoppers for acquisition.

Figure 8. Retailer's Revenue Loss by Percentile of Shoppers Experiencing Failure



Note: CATE = Conditional Average Treatment Effect.

Recovery Strategies

The finding that app failures result in reduced purchases across channels suggests that managers should develop interventions and recovery strategies to mitigate the negative effects of app failures not just in the mobile channel, but also in other channels, in particular, the offline channel. Thus, seamlessly integrating data from a mobile app with data from its stores and websites can help a multichannel retailer build continuity in shoppers' experiences.

Immediately after a shopper experiences an app failure, the manager of the app should provide gentle nudges and even incentives for the shopper to complete an abandoned transaction on the app. Typically, a manager may need to provide these nudges and incentives through other communication channels such as email, phone call, or face-to-face chat. These nudges are similar in spirit and execution to those from firms like Fitbit and Amazon, who remind customers through email to reconnect when they disconnect their watch and smart speaker, respectively. If the store is a dominant channel for the retailer, the retailer should use its store associates to reassure or incentivize shoppers. In some cases, managers can even offer incentives in other channels to complete a transaction disrupted by an app failure.

Because diminished purchases after failure result from reduced engagement, managers should aim to enhance engagement after a systemwide failure. Once a failure is restored, managers could induce shoppers to use the app more through gamification features in the app or providing enhanced loyalty points for logging back into the app.

The finding that app failure can enhance spending for shoppers experiencing the failure close to the store offers useful cross-selling opportunities for the retailer. After a systemwide failure is resolved, retailers can proactively promote in the store nearest to each failure-experiencing shopper products based on the shoppers' purchase history.

Managers should mitigate the negative effects of app failures for the most sensitive shoppers first. They should proactively identify failure-sensitive shoppers and design preemptive strategies to mitigate any adverse effects. We find that first-time digital channel users for a retailer and shoppers with weaker relationship with the provider are more sensitive to failures. Thus, firms should address such shoppers for recovery after a careful cost-benefit analysis. This is important because apps serve as a gateway for future purchases for these shoppers.

Finally, our analysis of heterogeneity in shoppers' sensitivity to app failures suggests that managers should selectively target shoppers for service recovery. Managers should satisfy first the shoppers with the highest values of CATE. Interventions targeted at the 51% of the shoppers who contribute to 70% of the revenue loss will likely lead to higher returns than other efforts.

Limitations

Our study has limitations that future research can address. First, we do not know the nature of each app failure, so we could not study the intensity of failure experience that could range from a temporary slowdown to a total shutdown. Second, our results are most informative for similar retailers that have a large brick-and-mortar presence with a small but growing online channel or app-induced checkouts and purchases. If data are available, future research could study app failures for primarily online retailers with an expanding offline presence (e.g., Bonobos, Warby Parker). Third, we do not have data on competing apps that shoppers may use. Additional research could study shoppers' switching behavior if data on competing apps are available. Fourth, our data contain relatively low number of purchases in the mobile channel. For better generalizability of the extent of spillover across channels, our analysis could be extended to contexts in which a substantial portion of purchase transactions are made within the mobile app. Fifth, we do not have data on purchases made through mobile apps vs. mobile web browsers. Examining the differences between these two mobile sub-channels is a fruitful avenue for future research. Finally, mobile apps may provide an effective means to resolve problems and recover from the adverse effects of service failures (Tucker and Yu 2018). Future research can collect data on service recovery strategies and shoppers' responses to them to identify the best mitigation tactics. Thus, the proposed preventive and service recovery strategies could be tested in ethically permissible situations if appropriate data can be gathered.

CHAPTER V

INTEGRATING THE TWO ESSAYS: A DISCUSSION

Mobile apps offer a unique capability to impact shoppers' omnichannel behaviors. Essays 1 and 2 of this dissertation, together, highlight the strategic importance of mobile apps for omnichannel retailers with a view toward providing powerful new insights for both theory and practice. While both essays address the important broad issue of leveraging mobile apps for engaging customers and impacting retailer revenues in multiple channels, independently they deal with specific objectives toward this end. Essay 1 highlights the *bright side* of retailers' branded mobile app introduction while Essay 2 uncovers a *dark side* of such app investments via app failures.

Specifically, Essay 1 quantifies the impact of mobile app introduction on online and offline purchases and product returns. In doing so, this essay allows retailers and researchers to understand the implications of mobile app introduction strategy for overall shopper purchases as well as purchases in specific channels. It is also among the first mobile marketing studies to examine product returns, and thereby net purchases. Overall, Essay 1 is a unique inquiry into the role mobile apps can play for large-scale retailers in enabling them to compete with online giants. It demonstrates that there is a 37% lift in net purchases in online and offline channels due to the launch of an app over 18 months for a large-scale retailer.

Essay 2 goes beyond the app introduction decision and examines a specific app usage experience by investigating failures in an app. Failure events are common and widespread in the technological setting that an app offers. However, it was not known prior to this essay whether failures in an app can cause shifts in shopping outcomes in other channels, both online and offline. Thus, Essay 2 complements and extends Essay 1 by going beyond the strategy for app introduction and further examining issues relating to managing and maintaining service delivery experiences through an app. Once a retailer has launched an app, what does it entail to create successful user experiences? In other words, what impact would negative app experiences have on shoppers within and beyond the digital channels? Essay 2 quantifies the impact of mobile app failures on online and offline purchases. In addition, it examines theoretical mechanisms underlying the effects across channels and in each channel, highlighting brand preference dilution and channel substitution as potential reasons. Importantly, it uncovers the heterogeneity among shoppers based on past relationship with the retailer. Overall, the results show a 7.1% decrease in purchases primarily in stores due to an app failure over two weeks.

Given the inter-dependent nature of inquiry in these two essays and yet a distinctive focus on specific mobile app related strategy issues, these essays offer a wide range of implications for theory and practice. These essays are among the first foray into research on mobile apps' impact on offline outcomes beyond the online channel. In this sense, they help connect shoppers' online and offline footprints and understand their behaviors in multiple channels. This opens up huge opportunity for future research into these and similar technologies. Research in cross-device integration is a natural next step, one that is extremely critical in the rising age of Internet of Things (IoT) and Artificial Intelligence. Finally, while these essays examine traditionally brick-and-mortar retailers, there is huge potential to examine similar issues for online-first retailers and understand key differences in strategic implications of mobile technologies, both their introduction and potential setbacks or failures, on different types of retailers and shoppers. From a practitioner perspective, together, these essays have a potential impact of $504 million in economic terms. In percentage terms, the effect range from 7% of revenues in Essay 2 to 37% in Essay 1.

In an increasingly mobile-first world, understanding the impact of introduction of new technologies as well as preempting pitfalls of such technologies once launched is of critical importance. By examining both mobile app introduction and mobile app failures, and quantifying their impact for retailers' key outcomes in multiple channels, this dissertation contributes to the nascent but growing research in mobile and omnichannel marketing. It uncovers both a *bright side* and a potentially *dark side* of retailers' branded mobile apps and leads the pathway for future research in interconnected devices and new technologies impacting the retail landscape.

CHAPTER VI

CONCLUSIONS

The objective of this dissertation is to examine the impact of mobile app introduction and mobile app failures on omnichannel retailers' revenues in multiple channels. Specifically, it examined the following questions: Do mobile apps influence shoppers' online and offline purchases and product returns? Further, do in-app experiences, such as app failures experienced by shoppers influence their purchases? By doing so, this dissertation expands the scope of mobile marketing literature by quantifying and explaining the impact of mobile app introduction and app failures on shopping across channels. Unlike prior studies that focus on only online purchases, I consider both online and in-store purchases, and product returns.

The key findings show that simply by launching a branded retail app, retailers can get a 37% lift in net purchases in online and offline channels due to the launch of an app over 18 months. However, apps have a dark side too, and can expose shoppers to negative failure occurrences. A single failure can result in 7.1% loss of revenues in offline channels over two weeks after the failure and the effects persist over several weeks. Therefore, while launching an app has a positive net effect on retailer revenues in multiple channels, a faulty or failing app can also cause significant damage. Essay 1 highlights the *bright side* of retailers' branded mobile app introduction while Essay 2 uncovers a *dark side* of such app investments via app failures. Together, the two essays suggest that retailers expanding into new mobile technologies must maximize the gains from launching a new branded app while minimizing the potential risks due to failure.

REFERENCES

Adamopoulos P, Tuzhilin A (2015) The business value of recommendations: A privacy-preserving econometric analysis. *Working Paper.*

Andreassen TW (1999) What drives customer loyalty with complaint resolution? *J. Service Res.* 1(4):324-332.

Andrews M, Goehring J, Hui S, Pancras J, Thornswood L (2016) Mobile promotions: A framework and research priorities. *J. Interactive Marketing* 34:15-24.

Andrews M, Luo X, Fang Z, Ghose A (2015) Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Sci.* 35(2):218-233.

Angrist JD, Pischke JS (2009) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, Princeton, NJ).

Altonji JG, Elder TE, Taber CR (2005) Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *J. Political Economy* 113(1):151-184.

Ansari A, Mela CF, Neslin SA (2008) Customer channel migration. *J. Marketing Res.* 45(1): 60-76.

App Annie (2018) 2017 Retrospective: A monumental year for the app economy. Retrieved from https://tinyurl.com/ycrlmq56

Ascarza E (2018) Retention futility: Targeting high-risk customers might be ineffective. *J. Marketing Res.* 55(1):80-98.

Athey S, Imbens G (2016) Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sci.* 113(27):7353-7360.

Athey S, Imbens G, Pham T, Wager S (2017) Estimating average treatment effects: Supplementary analyses and remaining challenges. *American Econom. Rev.* 107(5):278-81.

Avery J, Steenburgh T, Deighton J, Caravella M (2012) Adding bricks to clicks: Predicting the patterns of cross-channel elasticities over time. *J. Marketing* 76(3):96-111.

Barrons (2018) Walmart: Can it meet its digital sales growth targets? Accessed May 20, 2019, https://tinyurl.com/y266gykm.

Bart Y, Stephen A T, Sarvary M (2014) Which products are best suited to mobile advertising? A field study of mobile display advertising effects on consumer attitudes and intentions. *J. Marketing Res.* 51(3):270-285.

Barwise P, Strong C (2002) Permission-based mobile advertising. *J. Interactive Marketing* 16(1):14-24.

BCG (2017) Revolution in the driver's seat: The road to autonomous vehicles. Retrieved from https://tinyurl.com/ycx69d7c

Belk R (2014) You are what you can access: Sharing and collaborative consumption online. *J. Bus. Res.* 67(8):1595-1600.

Bell D, Gallino S, Moreno A (2017) Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Sci.* 64(4):1629-1651.

Bellman S, Potter RF, Treleaven-Hassard S, Robinson JA, Varan D (2011) The effectiveness of branded mobile phone apps. *J. Interactive Marketing* 25(4):191–200.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust difference-in-difference estimates? *Quart. J. Econom.* 119(1):249–275.

Bitner MJ, Booms BH, Tetreault MS (1990) The service encounter: Diagnosing favorable and unfavorable incidents. *J. Marketing* 71-84.

Blancco (2016) The state of mobile device performance and health: Q2. Accessed April 20, 2018, https://tinyurl.com/y9kd6bmz

Bolton RN (1998) A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. *Marketing Sci.* 17(1):45-65.

Boyd ED, Kannan PK, Slotegraaf RJ (2019) Branded apps and their impact on firm value: A design perspective. *J. Marketing Res.* 56(1):76–88.

Breiman L (2001) Random forests. *Machine learning* 45(1):5-32.

Bronnenberg BJ, Dubé JPH, Gentzkow M (2012) The evolution of brand preferences: Evidence from consumer migration. *American Economic Rev.* 102(6):2472-2508.

Brynjolfsson E, Hu YJ, Rahman MS (2013) Competing in the age of omnichannel retailing. *MIT Sloan Management Rev.* 54(4):23.

Chandrashekaran M, Rotte K, Tax SS, Grewal R (2007) Satisfaction strength and customer loyalty. *J. Marketing Res.* 44(1):153-163.

Chintagunta PK, Chu J, Cebollada J (2012) Quantifying transaction costs in online/off-line grocery channel choice. *Marketing Sci.* 31(1):96-114.

Cleeren K, Dekimpe MG, Helsen K (2008) Weathering product-harm crises. *J. Academy of Marketing Sci.* 36(2):262-270.

Cleeren K, Van Heerde HJ, Dekimpe MG (2013) Rising from the ashes: How brands and categories can overcome product-harm crises. *J. Marketing* 77(2):58-77.

CNBC (2018a) Augmented reality could be an $8 billion ‘revenue opportunity’ for Apple, analyst says. Retrieved from https://tinyurl.com/y89wl59v

CNBC (2018b) The cutting edge of customer experience is… apps? *CNBC* (July 6), https://tinyurl.com/y8o4ukuv

comScore (2016) The 2016 US mobile app report. *ComScore* (September 13), http://tinyurl.com/j43tjyw

comScore. (2017) Mobile hierarchy of needs. Retrieved from www.comscore.com/mobile-hierarchy

Danaher PJ, Smith MS, Ranasinghe K, Danaher TS (2015) Where, when, and how long: Factors that influence the redemption of mobile phone coupons. *J. Marketing Res*. 52(5):710–725.

Darani M, Shankar V (2018) Topic hidden Markov modeling: A new machine learning approach to make dynamic purchase predictions. Working Paper, Texas A&M University.

Dickinger A, Kleijnen M (2008) Coupons going wireless: Determinants of consumer intentions to redeem mobile coupons. *J. Interactive Marketing* 22(3):23-39.

Dimensional Research (2015) Failing to meet mobile app user expectations: A mobile app user survey. Accessed September 24, 2017, https://tinyurl.com/ybbnffcc

Dotzel T, Shankar V, Berry LL (2013) Service innovativeness and firm value. *J. Marketing Res.* 50(2):259-276.

Dukes, A, Zhu Y (2019) Why customer service frustrates consumers: Exploiting hassle costs by a tiered organizational structure. *Marketing Sci.* 38(3):500-515.

Dubé JP, Fang Z, Fong N, Luo X (2017) Competitive price targeting with smartphone coupons. *Marketing Sci.* 36(6):944-975.

Einav L, Levin J, Popov I, Sundaresan N (2014) Growth, adoption, and use of mobile e-commerce. *American Economic Rev.* 104(5):489-494.

eMarketer (2017), eMarketer Unveils New Estimates for Mobile App Usage. *Emarketer* (April 11), https://tinyurl.com/ydyz4hsm

eMarketer. (2018) eMarketer: Mobile ad spending to surpass TV in 2018. Retrieved from https://tinyurl.com/yckmugq5

Engadget. (2018) Amazon's AR shopping tool is now available on Android. Retrieved from https://tinyurl.com/y7v7b27m

Faccio M, McConnell JJ (2018) Death by Pokémon GO: The economic and human cost of using apps while driving. National Bureau of Economic Research (No. w24308).

Fong NM, Fang Z, Luo X (2015) Geo-conquesting: Competitive locational targeting of mobile promotions. *J. Marketing Res.* 52(5):726–735.

Forbes (2015) How mobile ordering can impact Starbucks' valuation. *Forbes* (October 30), http://tinyurl.com/zrekrwp

Forbes (2016) How leading brands are winning the ‘direct-to-customer’ conversation. Retrieved from https://tinyurl.com/yd242dk5

Forbes (2017) 19 Crimes wine is an amazing example of adult targeted augmented reality. Retrieved from https://tinyurl.com/yah8e23t

Forbes (2018) Amazon and Marriott want to modernize hotel experiences. Retrieved from https://tinyurl.com/y8k5lhok/

Forbes LP (2008) When something goes wrong and no one is around: Non-internet self-service technology failure and recovery. *J. Services Marketing* 22(4):316-327.

Forbes LP, Kelley SW, Hoffman KD (2005) Typologies of e-commerce retail failures and recovery strategies. *J. Services Marketing* 19(5):280-292.

Gelb (2017) Branding, marketing, and the impact of self-driving cars. Retrieved from https://tinyurl.com/ydehc3p7

Ghose A, Kwon H, Lee D, Oh W (2019) Seizing the commuting moment: Contextual targeting based on mobile transportation apps. *Inform. Systems Res.* 30(1):154-174.

Ghose A (2017) TAP: Unlocking the Mobile Economy. Massachusetts, USA: MIT Press.

Ghose A, Goldfarb A, Han S P (2012) How is the mobile Internet different? Search costs and local activities. Information Systems Research, 24(3):613-631.

Ghose A, Guo X, Li B (2017) Empowering patients using smart mobile health platforms: Evidence from a randomized field experiment. *Working Paper, New York University, NY.*

Ghose A, Han S, Park S (2013) Analyzing the interdependence between Web and mobile advertising: A randomized field experiment. *Leonard N. Stern School of Business New York University Working Paper.*

Ghose A, Kwon H E, Lee D, Oh W (2018) Seizing the commuting moment: Contextual targeting based on mobile transportation apps. *Information Systems Res.* 30(1):154-174.

Ghose A, Li B, Liu S (2016) Mobile targeting using customer trajectory patterns. *Working Paper, New York University, NY.*

Grewal D, Ahlbom CP, Beitelspacher L, Noble SM, Nordfält J (2018) In-store mobile phone use and customer shopping behavior: Evidence from the field. *J. Marketing* 82(4):102-126.

Grewal D, Ahlbom CP, Noble SM, Shankar V, Narang U, Nordfält J (2018) Mobile integrated kiosks: How communication content increases spending. *Working Paper, Babson College, MA.*

Grewal D, Bart Y, Spann M, Zubcsek PP (2016) Mobile advertising: A framework and research agenda. *J. Interactive Marketing* 34:3-14.

Gijsenberg MJ, Van Heerde HJ, Verhoef PC (2015) Losses loom longer than gains: Modeling the impact of service crises on perceived service quality over time. *J. Marketing Res.* 52(5):642-656.

Gill M, Sridhar S, Grewal R (2017) Return on engagement initiatives: A study of a business-to-business mobile app. *J. Marketing* 81(4):45-66.

Goldfarb A, Tucker C (2011) Privacy regulation and online advertising. *Management Sci.* 57(1):57-71.

Goldfarb A, Tucker C (2014) Conducting research with quasi-experiments: A guide for marketers. *SSRN.*

Gomez MI, McLaughlin EW, Wittink DR (2004) Customer satisfaction and retail sales performance: An empirical investigation. *J. Retailing* 80(4):265-278.

Goodman PS, Fichman M, Lerch FJ, Snyder PR (1995) Customer-firm relationships, involvement, and customer satisfaction. *Academy of Management Journal* 38(5):1310-1324.

Google M/A/R/C Study (2013) Mobile in-store research: How in-store shoppers are using mobile devices. Google M/A/R/C. Accessed September 24, 2017, http://tinyurl.com/gr7ghpn

Google (2018) Mobile has changed search intent and how people get things done: New consumer behavior data. Retrieved from https://www.thinkwithgoogle.com/consumer-insights/mobile-search-consumer-behavior-data

Guo S, Fraser MW (2014) *Propensity Score Analysis: Statistical Methods and Applications* (Sage Publications).

Guo T, Sriram S, Manchanda P (2018) The effect of information disclosure on industry payments to physicians. *Working Paper.* Available at SSRN: https://ssrn.com/abstract=3064769.

Halbheer D, Gärtner DL, Gerstner E, Koenigsberg O (2018) Optimizing service failure and damage control. *Int. J. Res. in Marketing* 35(1):100-115.

Hansen N, Kupfer AK, Hennig-Thurau T (2018) Brand crises in the digital age: The short-and long-term effects of social media firestorms on consumers and brands. *Int. J. Res. in Marketing* 35(4):557-574.

Heckman J (1979) Sample selection bias as a specification error. *Econometrica* 47(1):153-161.

Heilman C M, Nakamoto K, Rao A G (2002) Pleasant surprises: Consumer response to unexpected in-store coupons. *J. Marketing Res.* 39(2):242-252.

Hess RL, Ganesan S, Klein NM (2003) Service failure and recovery: The impact of relationship factors on customer satisfaction. *J. Academy of Marketing Sci.* 31(2):127.

Hoffman KD, Bateson JE (1997) *Essentials of Services Marketing*. The Dryden Pres, USA.

Huang Q, Nijs VR, Hansen K, Anderson ET (2012) Wal-Mart's impact on supplier profits. *J. Marketing Res.* 49(2):131–143.

Hui S K, Inman J J, Huang Y, Suher J (2013) The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *J. Marketing* 77(2):1-16.

Hung SY, Ku CY, Chang CM (2003) Critical factors of WAP services adoption: An empirical study. *Electronic Commerce Research and Applications* 2(1):42-60.

Hyken S (2017) Walgreens: At the corner of technology and a better customer experience (CX). *Forbes* (July 8), http://tinyurl.com/y96uddk4

IAB. (2016) Digital ad spend reaches an all-time high of $88 billion in 2017, with mobile upswing unabated, accounting for 57% of revenue. Retrieved from https://tinyurl.com/y9frnohx

Jacobi L, Sovinsky M (2016) Marijuana on main street? Estimating demand in markets with limited access. *American Economic Rev.* 106(8):2009-45.

Jacobs, B.J.D., Donkers, B., Fok, D. (2016) Model-based purchase predictions for large assortments. Marketing Science, 35(3), 389-404.

Kesler, R., Kummer, M., Schulte, P. (2018) Mobile applications and access to private data: The supply side of the Android ecosystem. *Working Paper.*

Kim SJ, Wang R J-H, Malthouse EC (2015) The effects of adopting and using a brand's mobile application on customers' subsequent purchase behavior. *J. Interactive Marketing* 31:28–41.

Knox G, Van Oest R (2014) Customer complaints and recovery effectiveness: A customer base approach. *J. Marketing* 78(5):42-57.

Knowledge @ Wharton. (2016) Why 'missed call' marketing has taken hold in India. Retrieved from http://knowledge.wharton.upenn.edu/article/missed-call-marketing-taken-hold-india/

Kone (2018) AI, more of a reality than you think. Retrieved from https://tinyurl.com/y2qoa6ar

Kumar V, Dogan OB, Lahiri A (2016) Engaging customers in the app world through smart analytics. *J. World Marketing Summit*, 2(1).

Leanplum. (2017) Personalize or bust? The impact on app engagement. Retrieved from https://tinyurl.com/y8vcj625

Lee J, Zhuang M, Kozlenkova I, Fang E (2016) The dark side of mobile channel expansion strategies. *MSI Working Paper.*

Li C, Luo X, Zhang C, Wang X (2017) Sunny, rainy, and cloudy with a chance of mobile promotion effectiveness. *Marketing Sci.* 36(5):762-779.

Liu Y, Shankar V (2015) The dynamic impact of product-harm crises on brand preference and advertising effectiveness: An empirical analysis of the automobile industry. *Management Sci.* 61(10):2514-2535.

Lundborg P, Plug E, Rasmussen AW (2017) Can women have children and a career? IV evidence from IVF treatments. *American Economic Rev.* 107(6):1611-37.

Luo X, Andrews M, Fang Z, Phang CW (2013) Mobile targeting. *Management Sci.* 60(7):1738-1756.

Ma L, Sun B, Kekre S (2015) The squeaky wheel gets the grease: An empirical analysis of customer voice and firm intervention on Twitter. *Marketing Sci.* 34(5):627-645.

Manchanda P, Packard G, Pattabhiramaiah A (2015) Social dollars: The economic impact of customer participation in a firm-sponsored online customer community. *Marketing Sci*. 34(3):367-387.

Marshall, M. (2018) TGI Friday's AI-powered marketing drives $150 million (with virtual bartenders and more), Venturebeat. Retrieved from https://tinyurl.com/yd5mzjbr

McCollough MA, Berry LL, Yadav MS (2000) An empirical investigation of customer satisfaction after service failure and recovery. *J. Service Res.* 3(2):121-137.

Meuter ML, Ostrom AL, Roundtree RI, Bitner MJ (2000) Self-service technologies: understanding customer satisfaction with technology-based service encounters. *J. Marketing* 64(3):50-64.

Muir D, Tsai Y (2016) Search costs and consideration set formation with UGC: Fixed v. mobile devices. *Working Paper.*

Narang U, Shankar V (2019a) Mobile app introduction and online and offline purchases and product returns. *Marketing Sci.* 38(5):756-772.

Narang U, Shankar V (2019b) Mobile marketing 2.0: State of the art and research agenda. *Marketing in a Digital World, Review of Marketing Res.* 16:97-119.

National Retail Federation (2018) Stores top 100 retailer, Accessed January 5, 2019, https://stores.org/stores-top-retailers-2018/

Neumann N, Tucker CE, Whitfield T (2019) How effective is black-box digital consumer profiling and audience delivery?: Evidence from field studies. *Marketing Sci*. Forthcoming.

Oliver RL (1980) A cognitive model of the antecedents and consequences of satisfaction decisions. *J. Marketing Res.* 460-469.

Parasuraman A, Zeithaml VA, Berry LL (1994) Reassessment of expectations as a comparison standard in measuring service quality: Implications for future research. *J. Marketing* 58:111-124.

Parasuraman A, Zeithaml VA, Berry LL (1988) SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *J. Retailing* 64(1):12–37.

Pauwels K, Neslin SA (2015) Building with bricks and mortar: The revenue impact of opening physical stores in a multichannel environment. *J. Retailing* 91(2):182-197.

Pew Research Center (2018) Mobile Fact Sheet, http://www.pewinternet.org/fact-sheet/mobile/.

Rafieian O, Yoganarasimhan H (2018) Optimal sequencing of mobile advertisements. Working Paper, University of Washington, Seattle, WA.

Retail Dive (2018) App Annie: Shopping app use jumped 54% in 2017. *Retail Dive* (January 12), https://tinyurl.com/ycgnyj8o

Rosenbaum PR (2005) Sensitivity analysis in observational studies. Everitt BS, Howell DC, eds. *Encyclopedia of Statistics in Behavioral Science* (Wiley, New York), 1809-1814.

Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55.

Rubin DB (2008) For objective causal inference, design trumps analysis. *The Ann. of Appl. Statis.* 2(3):808–840.

Sandler, E. (2018) How L'Occitane is using AI to improve customer experiences online. Glossy. Retrieved from https://tinyurl.com/yazasd48

Schmittlein DC, Morrison DG, Colombo R (1987) Counting your customers: Who are they and what will they do next? *Management Sci.* 33(1):1-24.

Shankar V (2012) Mobile marketing strategy. In V. Shankar G. S. Carpenter (Eds.), Handbook of Marketing Strategy. Massachusetts, USA: Edward Elgar Publishing, 217-230.

Shankar V (2018) How artificial intelligence (AI) is reshaping retailing. *J. Retailing* 94(4):6-11.
Shankar V, Balasubramanian S (2009) Mobile marketing: A synthesis and prognosis. *J. Interactive Marketing* 23(2):118–129.

Shankar V, Hollinger M (2007) Online and mobile advertising: Current scenario, emerging trends, and future directions. *Marketing Science Institute* 31(3):206-207.

Shankar V, Kleijnen M, Ramanathan S, Rizley R, Holland S, Morrissey S (2016) Mobile shopper marketing: Key issues, current insights, and future research avenues. *J. Interactive Marketing* 34:37-48.

Shankar V, O Driscoll T, Reibstein D (2003) Rational exuberance: The wireless industry's Killer 'B'. Strategy + Business, 31 (Summer), 68-77.

Shankar V, Venkatesh A, Hofacker C, Naik P (2010) Mobile marketing in the retailing environment: Current insights and future research avenues. *J. Interactive Marketing* 24(2):111-120.

Shi S, Kalyanam K, Wedel M (2017) What does agile and lean mean for customers? An analysis of mobile app crashes. *Working Paper.*

Silverman BW (1986) *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, London).

Sims G (2015) Google Play store vs the Apple App store: By the numbers. *Android Authority* (April 20), http://tinyurl.com/zp4ufdq

Smith AK, Bolton RN (1998) An experimental investigation of customer reactions to service failure and recovery encounters: Paradox or peril? *J. Service Res.* 1(1):65-81.

Statista (2015) Share of console gamers in the United States who play online and offline games as of April 2015, by platform. *Statista,* https://tinyurl.com/ydf7rflp

Statista. (2018) Number of smartphone users in the United States from 2010 to 2022 (in millions) Retrieved from https://tinyurl.com/j2mzjqh

Sudhir K, Talukdar D (2015) The "Peter Pan syndrome" in emerging markets: The productivity-transparency trade-off in IT adoption. *Marketing Sci.* 34(4):500-521.

Tax SS, Brown SW, Chandrashekaran M (1998) Customer evaluations of service complaint experiences: Implications for relationship marketing. *J. Marketing.* 62(2):60-76.

TechCrunch. (2017) U.S. consumers now spend 5 hours per day on mobile devices. Retrieved from https://tinyurl.com/y7dvmx7b

Tirunillai S, Tellis G (2017) Does offline TV advertising affect online chatter? Quasi-experimental analysis using synthetic control. *Marketing Sci.* 36(6):862-878.

Tucker CE, Yu S (2018) Does it lead to more equal treatment? An empirical study of the effect of smartphone use on customer complaint resolution. *Working Paper.* Massachusetts Institute of Technology MA.

Venture Beat (2016) Video games will become a $99.6B industry this year as mobile overtakes consoles and PCs. *Venture Beat* (April 21), http://tinyurl.com/y7favjyc

Wager S, Athey S (2018) Estimation and inference of heterogeneous treatment effects using random forests. *J. the American Statistical Association* 113(523):1228-1242.

Wakabayashi, D. (2018) California passes sweeping law to protect online privacy. *The New York Times*. Retrieved from https://tinyurl.com/y7ojx39l

Waldrop MM (2015) No drivers required. *Nature* 518(7537):20.

Wang K, Goldfarb A (2017) Can offline stores drive online sales? *J. Marketing Res.* 54(5):706-719.

Wang RJ-H, Malthouse EC, Krishnamurthi L (2015) On the go: How mobile shopping affects customer purchase behavior. *J. Retailing* 91(2):217–234.

Wangenheim FV, Bayón T (2007) Behavioral consequences of overbooking service capacity. *J. Marketing* 71(4):36–47.

Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge).

Xu K, Chan J, Ghose A, Han SP (2016) Battle of the channels: The impact of tablets on digital commerce. *Management Sci.* 63(5):1469-1492.

Xu J, Forman C, Kim JB, Ittersum KV (2014) News media channels: Complements or substitutes? Evidence from mobile phone usage. *J. Marketing* 78(4):97–112.

Zubcsek PP, Katona Z, Sarvary M (2017) Predicting mobile advertising response using consumer colocation networks. *J. Marketing* 81(4):109-126.

APPENDIX A

APPENDIX FOR CHAPTER III

Table A1. Table 10 for an Alternative Sample of App Adopters

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
(App Adopters X Post App Introduction)	0.439*** (0.011)	0.873*** (0.025)	24.807*** (0.66)	0.062*** (0.003)	0.084*** (0.005)	2.284*** (0.167)	22.523*** (0.595)
App Adopters	-0.129*** (0.013)	-0.288*** (0.029)	-6.957*** (0.712)	-0.019*** (0.004)	-0.026*** (0.006)	-0.563** (0.171)	-6.394*** (0.644)
IMR	-1.183*** (0.022)	-2.314*** (0.051)	-61.778*** (1.138)	-0.208*** (0.009)	-0.289*** (0.012)	-5.973*** (0.316)	-55.805*** (1.003)
Intercept	1.586*** (0.025)	3.052*** (0.059)	76.588*** (1.302)	0.253*** (0.009)	0.342*** (0.013)	7.999*** (0.362)	68.589*** (1.147)
Number of Observations	3,600,000	3,600,000	3,600,000	3,600,000	3,600,000	3,600,000	3,600,000
R squared	0.075	0.056	0.054	0.023	0.019	0.005	0.057

**Notes:* Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value. IMR = inverse Mills ratio from the selection model.

Table A2. Table 11 for an Alternative Sample of App Adopters

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
(App Adopters X Post App Introduction)	0.375*** (0.011)	0.759*** (0.025)	21.318*** (0.636)	0.065*** (0.001)	0.115*** (0.002)	3.489*** (0.112)
App Adopters	-0.125*** (0.013)	-0.283*** (0.029)	-6.755*** (0.692)	-0.004*** (0.001)	-0.005** (0.002)	-0.202* (0.091)
IMR	-1.147*** (0.022)	-2.258*** (0.05)	-59.703*** (1.1)	-0.036*** (0.002)	-0.056*** (0.002)	-2.075*** (0.111)
Intercept	1.544*** (0.025)	2.993*** (0.058)	74.483*** (1.267)	0.042*** (0.002)	0.059*** (0.003)	2.104*** (0.155)
Number of Observations	3,600,000	3,600,000	3,600,000	3,600,000	3,600,000	3,600,000
R squared	0.076	0.057	0.056	0.078	0.056	0.026

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. IMR = inverse Mills ratio from the selection model.

Table A3. App Feature Categorization

Feature bucket	Descriptors
Banner	Home-screen banner
Product	Product check availability, product details, product scan barcode, product search catalog, product view screenshots, product view video
Store	Store locator, call store, check-in complete, check-in start, get directions, store details
Shop	Checkout tap, add to cart
Offers	Active offers details, active offers list
Rewards	Rewards landing, activate tap, reward redeem complete, reward redeem start, reward button tap, rewards search, reward account details
Notifications	Message center tap, push notification acted on, notifications registered

Table A4. Logit Model for Propensity Score Matching

Variable	Coeff. (Std. Err.)
Intercept	2.364***(0.057)
Age	-0.080***(0.003)
Gender	0.038***(0.011)
Recency	-0.005***(0.000)
Loyalty program level	-0.212***(0.018)
No. of stores in zip code	0.248***(0.025)
Area population	0.000***(0.000)
Distance to nearest store	0.049***(0.001)
No. of competitor stores	0.003(0.014)
Age squared	0.001***(0.000)
Number of stores squared	-0.075***(0.010)
Distance to nearest store squared	0.000***(0.000)
Past frequency of purchases – online	0.039(0.020)
Past quantity of purchases – online	0.042**(0.014)
Past value of purchases – online	-0.001***(0.000)
Past discounts used	0.451***(0.021)
Past quantity purchased per order	-0.636***(0.012)

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Additional covariates, frequency, quantity, and value of purchases across channels and returns for each month in the pre period (108 variables), not reported to save space.

Table A5. Match Balance after Propensity Score Matching

	Means treated	Means control (before matching)	Means control (after matching)	Mean diff.
Propensity score	0.6251	0.3299	0.6251	100.0000
Age	30.5808	33.2083	30.6556	97.1514
Gender	0.7670	0.5938	0.7898	86.8397
Recency	71.292	301.2206	69.8841	99.3877
Loyalty program level	0.1705	0.1544	0.1718	92.2993
No. of stores in zip code	0.5709	0.5417	0.5860	48.5208
Area population	31762.9866	31151.388	32028.3817	56.6063
Distance to nearest store	3.7106	4.0606	3.5740	60.9706
No. of competitor stores	0.3041	0.2826	0.3180	35.116
Age squared	1031.1169	1233.5082	1041.4871	94.8762
Number of stores squared	0.8453	0.8064	0.8676	42.5847
Distance to nearest store squared	67.2289	67.1945	56.8629	-29991.94
Past frequency of purchases – online	0.3461	0.1140	0.3559	95.8145
Past quantity of purchases – online	0.5011	0.1635	0.5013	99.9254
Past value of purchases – online	25.3846	8.9046	28.2858	82.3953
Past discounts used	0.8974	0.5245	0.9084	97.0614
Past quantity purchased per order	1.8775	1.4971	1.8652	96.7751

Note: Additional covariates for matching, frequency, quantity, and value of purchases across channels and returns for each month in the pre period (108 variables), not reported to save space.

Table A6. Estimates by Month for Value of Purchases Across Channels

Variable	Coeff. (Std. Err.)
App Adopters X Month 2	1.739 (1.701)
App Adopters X Month 3	-0.288 (2.075)
App Adopters X Month 4	0.632 (1.492)
App Adopters X Month 5	-0.143 (1.617)
App Adopters X Month 6	0.962 (1.799)
App Adopters X Month 7	0.375 (1.771)
App Adopters X Month 8	0.021 (1.954)
App Adopters X Month 9	1.52 (1.73)
App Adopters X Month 10	-0.72 (2.17)
App Adopters X Month 11	5.245 (4.772)
App Adopters X Month 12	2.776 (2.489)
App Adopters X Month 13	-1.27 (2.036)
App Adopters X Month 14	1.932 (2.603)
App Adopters X Month 15	1.208 (2.983)
App Adopters X Month 16	3.152 (2.066)
App Adopters X Month 17	1.332 (2.164)
App Adopters X Month 18	2.005 (2.007)
-- *App Introduction* --	
App Adopters X Month 19	10.021*** (1.653)
App Adopters X Month 20	14.357*** (1.765)
App Adopters X Month 21	22.307*** (1.943)
App Adopters X Month 22	22.969*** (2.191)
App Adopters X Month 23	41.52*** (2.807)
App Adopters X Month 24	31.205*** (2.215)
App Adopters X Month 25	18.521*** (1.619)
App Adopters X Month 26	31.04*** (2.217)
App Adopters X Month 27	25.84*** (1.984)
App Adopters X Month 28	19.22*** (1.933)
App Adopters X Month 29	16.549*** (1.652)
App Adopters X Month 30	25.421*** (2.058)
App Adopters X Month 31	18.172*** (1.861)
App Adopters X Month 32	17.559*** (1.563)
App Adopters X Month 33	24.377*** (1.83)
App Adopters X Month 34	26.347*** (2.344)
App Adopters X Month 35	45.749*** (2.56)
App Adopters X Month 36	27.888*** (2.129)
App Adopter	-0.638 (1.519)
Intercept	38.879*** (1.476)
Number of observations	4,001,760
R-squared	0.0384

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included but not reported to save space; this table offers a formal test of Figure 2 and shows that the estimates were not significantly different from zero in the pre-app introduction period (prior to month 19, July 2014) but are significantly higher in each month post app launch; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table A7. Table 11 Alternative Model with Future App Adopters as Control

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
(App Adopters X Post App Introduction)	0.245***	0.458***	14.419***	0.017***	0.024***	1.115***
	(0.014)	(0.031)	(1.171)	(0.002)	(0.003)	(0.186)
App Adopters	0.007	0.014	0.419	0	0	-0.073
	(0.022)	(0.047)	(1.451)	(0.002)	(0.003)	(0.178)
IMR	-3.202***	-5.896***	-170.31***	-0.046***	-0.063***	-3.372***
	(0.077)	(0.147)	(4.079)	(0.004)	(0.007)	(0.326)
Intercept	9.607***	17.717***	498.998***	0.138***	0.194***	9.895***
	(0.203)	(0.391)	(10.824)	(0.011)	(0.018)	(0.883)
Number of Observations	1,388,772	1,388,772	1,388,772	1,388,772	1,388,772	1,388,772
R squared	0.0179	0.0125	0.0384	0.0026	0.0017	0.0025

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; in this method, matched future app adopters from April-December 2015 are used as controls for app adopters from July 2014-March 2015; their outcomes pre- and post- 9 months are compared; *** p < 0.001, ** p < 0.01, * p < 0.05; IMR = inverse Mills ratio from the selection model.

Table A8. Table 10 with Inverse Mills Ratio (IMR) from Selection Model without Matching

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
(App Adopters X Post App Introduction)	0.201***	0.353***	12.865***	0.025***	0.030***	1.170***	11.695***
	(0.005)	(0.011)	(0.293)	(0.001)	(0.002)	(0.071)	(0.264)
App Adopters	0.315***	0.599***	16.378***	0.055***	0.076***	1.596***	14.782***
	(0.004)	(0.009)	(0.239)	(0.001)	(0.002)	(0.061)	(0.217)
IMR	-0.818***	-1.515***	-42.888***	-0.124***	-0.167***	-3.656***	-39.232***
	(0.005)	(0.011)	(0.283)	(0.001)	(0.002)	(0.065)	(0.253)
Intercept	1.257***	2.295***	59.561***	0.179***	0.236***	5.638***	53.923***
	(0.007)	(0.015)	(0.371)	(0.002)	(0.003)	(0.103)	(0.334)
Number of Observations	4,274,784	4,274,784	4,274,784	4,274,784	4,274,784	4,274,784	4,274,784
R squared	0.168	0.121	0.104	0.044	0.033	0.011	0.106

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value. IMR = inverse Mills ratio from the selection model.

Table A9. Table 11 with Inverse Mills Ratio (IMR) from Selection Model without Matching

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
(App Adopters X Post App Introduction)	0.187***	0.331***	12.208***	0.014***	0.021***	0.656***
	(0.005)	(0.010)	(0.287)	(0.000)	(0.001)	(0.038)
App Adopters	0.310***	0.593***	16.046***	0.004***	0.006***	0.332***
	(0.004)	(0.009)	(0.235)	(0.000)	(0.001)	(0.031)
IMR	-0.803***	-1.494***	-41.913***	-0.015***	-0.021***	-0.975***
	(0.005)	(0.011)	(0.277)	(0.000)	(0.001)	(0.022)
Intercept	1.234***	2.262***	58.401***	0.023***	0.033***	1.16***
	(0.007)	(0.015)	(0.365)	(0.001)	(0.001)	(0.034)
Number of Observations	4,274,784	4,274,784	4,274,784	4,274,784	4,274,784	4,274,784
R squared	0.166	0.121	0.103	0.008	0.004	0.004

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. IMR = inverse Mills ratio from the selection model.

Table A10. Table 10 Robustness to Poisson Specification for Count-based Outcomes

Variable	Frequency of purchases	Quantity of purchases	Frequency of returns	Quantity of returns
(App Adopters X Post App Introduction)	0.340***	0.354***	0.374***	0.388***
	(0.002)	(0.001)	(0.005)	(0.004)
App Adopters	0.011***	0.019***	0.045***	0.051***
	(0.001)	(0.001)	(0.003)	(0.003)
Intercept	0.180***	0.795***	-1.776***	-1.486***
	(0.001)	(0.001)	(0.002)	(0.002)
Number of Observations	4,001,760	4,001,760	4,001,760	4,001,760
Log pseudo-likelihood	-7,121,318	-12,440,238	-2,080,736	-2,741,563

Notes: Robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table A11. Table 11 Robustness to Poisson Specification for Count-based Outcomes

	(1) Offline		(2) Online	
Variable	Frequency of purchases	Quantity of purchases	Frequency of purchases	Quantity of purchases
(App Adopters X Post App Introduction)	0.332*** (0.002)	0.349*** (0.001)	0.692*** (0.014)	0.712*** (0.012)
App Adopters	0.012*** (0.001)	0.019*** (0.001)	-0.028*** (0.010)	-0.001* (0.008)
Intercept	0.163*** (0.001)	0.782*** (0.001)	-3.924*** (0.007)	-3.581*** (0.006)
Number of Observations	4,001,760	4,001,760	4,001,760	4,001,760
Log pseudo-likelihood	-7,039,356	-12,325,174	-444,241	-648,911

Notes: Robust standard errors in parentheses; *** p < 0.001, ** p < 0.01, * p < 0.05.

Table A12. Table 10 Robustness to App Adoption Date as Cut-off for Pre- and Post- Periods

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
(App Adopters X Post Adoption)	0.322*** (0.012)	0.628*** (0.026)	17.979*** (0.744)	0.049*** (0.004)	0.066*** (0.007)	1.713*** (0.232)	16.267*** (0.664)
App Adopters	0.126*** (0.021)	0.251*** (0.038)	7.436*** (1.073)	0.025** (0.007)	0.035** (0.012)	0.738* (0.349)	6.698*** (0.91)
IMR	-2.446*** (0.044)	-4.504*** (0.082)	-122.227*** (2.083)	-0.380*** (0.016)	-0.508*** (0.025)	-11.23*** (0.623)	-110.99*** (1.795)
Intercept	7.094*** (0.124)	13.024*** (0.230)	346.955*** (5.922)	1.086*** (0.045)	1.445*** (0.07)	32.431*** (1.802)	314.524*** (5.099)
Number of Observations	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760
R squared	0.005	0.004	0.005	0.011	0.009	0.004	0.055

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value. IMR = inverse Mills ratio from the selection model.

Table A13. Table 11 Robustness to App Adoption Date as Cut-off for Pre- and Post- Periods

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
(App Adopters X Post Adoption	0.307*** (0.012)	0.605*** (0.026)	17.112*** (0.733)	0.014*** (0.001)	0.023*** (0.002)	0.868*** (0.109)
App Adopters)	0.122*** (0.021)	0.245*** (0.038)	7.294*** (1.047)	0.004*** (0.001)	0.006*** (0.002)	0.142 (0.104)
IMR	-2.411*** (0.044)	-4.454*** (0.081)	-119.92*** (2.049)	-0.035*** (0.001)	-0.05*** (0.003)	-2.306*** (0.122)
Intercept	6.991*** (0.123)	12.878*** (0.228)	340.55*** (5.829)	0.103*** (0.004)	0.146*** (0.008)	6.41*** (0.34)
Number of Observations	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760
R squared	0.049	0.037	0.051	0.003	0.002	0.003

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. IMR = inverse Mills ratio from the selection model.

Table A14. Table 10 Treatment Effect Coefficient for Different Time Periods

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
3 months	0.245*** (0.017)	0.470*** (0.038)	13.399*** (1.202)	0.039*** (0.007)	0.051*** (0.01)	0.548 (0.356)	12.851*** (1.09)
6 months	0.345*** (0.015)	0.656*** (0.031)	22.337*** (1.068)	0.055*** (0.006)	0.079*** (0.01)	2.039*** (0.331)	20.298*** (0.963)
9 months	0.359*** (0.013)	0.676*** (0.028)	22.458*** (1.032)	0.054*** (0.005)	0.073*** (0.009)	2.017*** (0.311)	20.441*** (0.885)
12 months	0.366*** (0.013)	0.681*** (0.027)	21.783*** (0.832)	0.055*** (0.004)	0.073*** (0.008)	2.016*** (0.231)	19.766*** (0.742)

Notes: Results are based on the models used for the main analysis pre-and post- app introduction date, that is, difference-in-differences for the 1:1 Nearest Neighbor matched sample with replacement. Note that between 3-6 months, effects increase strongly by (22.34 – 13.40) $8.94 whereas in later months the difference diminishes; robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value.

Table A15. Table 11 Treatment Effect Coefficient for Different Time Periods

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
3 months	0.238***	0.463***	13.137***	0.006***	0.007*	0.272*
	(0.017)	(0.038)	(1.188)	(0.002)	(0.003)	(0.117)
6 months	0.333***	0.638***	21.561***	0.014***	0.019***	0.851***
	(0.015)	(0.031)	(1.056)	(0.001)	(0.002)	(0.164)
9 months	0.345***	0.655***	21.389***	0.016***	0.023***	1.152***
	(0.013)	(0.028)	(1.024)	(0.001)	(0.002)	(0.193)
12 months	0.351***	0.659***	20.778***	0.017***	0.024***	1.087***
	(0.013)	(0.026)	(0.819)	(0.001)	(0.002)	(0.152)

Notes: Results are based on the models used for the main analysis pre-and post- app introduction date, that is, difference-in-differences for the 1:1 Nearest Neighbor matched sample with replacement. Between 3-6 months, effects increase strongly by \$8.42 (\$21.56–\$13.14) in-store and \$0.58 (\$0.85-\$0.27) online, whereas in later months, the difference diminishes; robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05.

Table A16. Table 10 Robustness to Alternative Matching Methods

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
Mahalanobis Metric	0.124***	0.231***	11.701***	0.018***	0.027***	1.490***	10.211***
	(0.006)	(0.013)	(0.398)	(0.002)	(0.002)	(0.106)	(0.354)
Nearest Neighbor (no replacement)	0.211***	0.372***	13.213***	0.027***	0.031***	1.211***	12.003***
	(0.005)	(0.011)	(0.299)	(0.001)	(0.002)	(0.073)	(0.269)
Caliper (0.27)	0.357***	0.663***	21.569***	0.05***	0.065***	2.052***	19.517***
	(0.006)	(0.012)	(0.362)	(0.001)	(0.002)	(0.088)	(0.33)
Caliper (0.05)	0.351***	0.651***	21.149***	0.049***	0.063***	1.951***	19.198***
	(0.006)	(0.012)	(0.363)	(0.001)	(0.002)	(0.087)	(0.331)

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; the numbers in parentheses for Caliper matching represent bandwidths of 0.27 and 0.05 times the standard deviation of propensity scores to test under stringent matching conditions as in Xu et. al (2016); *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value.

Table A17. Table 11 Robustness to Alternative Matching Methods

Variable	(1) Offline Frequency of purchases	(1) Offline Quantity of purchases	(1) Offline Value of purchases	(2) Online Frequency of purchases	(2) Online Quantity of purchases	(2) Online Value of purchases
Mahalanobis	0.112***	0.212***	10.982***	0.013***	0.019***	0.72***
	(0.006)	(0.012)	(0.39)	(0.001)	(0.001)	(0.059)
Nearest Neighbor (no replacement)	0.197***	0.351***	12.542***	0.014***	0.021***	0.672***
	(0.005)	(0.01)	(0.293)	(0)	(0.001)	(0.039)
Caliper (0.27)	0.344***	0.642***	20.798***	0.013***	0.021***	0.771***
	(0.006)	(0.012)	(0.354)	(0.001)	(0.001)	(0.049)
Caliper (0.05)	0.338***	0.63***	20.358***	0.013***	0.021***	0.791***
	(0.006)	(0.012)	(0.355)	(0.001)	(0.001)	(0.049)

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; the numbers in parentheses for Caliper matching represent bandwidths of 0.27 and 0.05 times the standard deviation of propensity scores to test under stringent matching conditions as in Xu et. al (2016); *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value.

Table A18. Table 10 Robustness to Outlier Spenders

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
(App Adopters X Post App Introduction)	0.236***	0.432***	14.996***	0.035***	0.044***	1.462***	13.534***
	(0.008)	(0.017)	(0.471)	(0.002)	(0.003)	(0.111)	(0.425)
App Adopters	0.468***	0.855***	23.544***	0.075***	0.102***	2.013***	21.531***
	(0.008)	(0.015)	(0.410)	(0.002)	(0.003)	(0.085)	(0.381)
IMR	-0.796***	-1.444***	-41.474***	-0.099***	-0.126***	-2.948***	-38.526***
	(0.009)	(0.016)	(0.445)	(0.002)	(0.002)	(0.061)	(0.416)
Intercept	2.479***	4.475***	123.672***	0.302***	0.384***	9.502***	114.17***
	(0.026)	(0.049)	(1.315)	(0.005)	(0.007)	(0.206)	(1.230)
Number of Observations	3,789,612	3,789,612	3,789,612	3,789,612	3,789,612	3,789,612	3,789,612
R squared	0.011	0.081	0.074	0.025	0.019	0.007	0.075

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value. IMR = inverse Mills ratio from the selection model.

Table A19. Table 11 Robustness to Outlier Spenders

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
(App Adopters X Post App Introduction)	0.223*** (0.008)	0.413*** (0.016)	14.430*** (0.460)	0.012*** (0.001)	0.019*** (0.001)	0.566*** (0.069)
App Adopters	0.459*** (0.008)	0.842*** (0.015)	22.939*** (0.403)	0.009*** (0.001)	0.012*** (0.001)	0.605*** (0.051)
IMR	-0.784*** (0.009)	-1.427*** (0.016)	-40.615*** (0.437)	-0.012*** (0.000)	-0.018*** (0.001)	-0.860*** (0.033)
Intercept	2.439*** (0.025)	4.420*** (0.048)	121.32*** (1.293)	0.039*** (0.002)	0.055*** (0.003)	2.356*** (0.100)
Number of Observations	3,789,612	3,789,612	3,789,612	3,789,612	3,789,612	3,789,612
R squared	0.011	0.081	0.072	0.006	0.003	0.004

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.05. IMR = inverse Mills ratio from the selection model.

Table A20. Table 10 Robustness to Pre-Period Deal Proneness

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
(App Adopters X Post App Introduction)	0.399*** (0.012)	0.749*** (0.024)	24.684*** (0.798)	0.049*** (0.002)	0.063*** (0.004)	1.910*** (0.153)	22.774*** (0.746)
App Adopters	-0.011 (0.01)	-0.016 (0.015)	-0.593 (0.498)	0.000 (0.002)	0.000 (0.002)	0.016 (0.088)	-0.609 (0.481)
IMR	-0.563*** (0.027)	-0.964*** (0.035)	-29.801*** (0.793)	-0.053*** (0.002)	-0.066*** (0.003)	-1.41*** (0.137)	-28.391*** (0.748)
Intercept	0.867*** (0.032)	1.454*** (0.042)	42.685*** (1.024)	0.089*** (0.004)	0.11*** (0.005)	2.768*** (0.236)	39.917*** (0.951)
Number of Observations	679,680	679,680	679,680	679,680	679,680	679,680	679,680
R squared	0.045	0.038	0.039	0.01	0.008	0.004	0.04

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included *** p < 0.001, ** p < 0.01, * p < 0.05. NMV = Net Monetary Value. IMR = inverse Mills ratio from the selection model.

Table A21. Table 11 Robustness to Pre-Period Deal Proneness

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
(App Adopters X Post App Introduction)	0.383*** (0.012)	0.727*** (0.023)	23.724*** (0.753)	0.017*** (0.002)	0.022*** (0.003)	0.960*** (0.208)
App Adopters	-0.009 (0.010)	-0.016 (0.015)	-0.470 (0.439)	-0.0020 (0.002)	0.000 (0.003)	-0.122 (0.233)
IMR	-0.54*** (0.027)	-0.932*** (0.035)	-28.018*** (0.731)	-0.023*** (0.004)	-0.032*** (0.006)	-1.783*** (0.340)
Intercept	0.832*** (0.032)	1.408*** (0.042)	40.346*** (0.946)	0.035*** (0.005)	0.046*** (0.007)	2.339*** (0.468)
Number of Observations	679,680	679,680	679,680	679,680	679,680	679,680
R squared	0.077	0.061	0.052	0.006	0.004	0.004

**Notes:* Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. IMR = inverse Mills ratio from the selection model.

Table A22. Top Three Most Frequently Used Features by Location and Time of App Use

Time to purchase (hours)	Distance from store (miles)			
	(0 to 1.30)	(1.31 to 3.11)	(3.12 to 7.23)	(7.24 or more)
(0 to 6.57)	Rewards landing	Product details	Product details	Product details
	Product details	Product details	Product details	Product details
	Rewards landing	Product search	Product search	Product search
(6.58 to 52.50)	Rewards landing	Product details	Product details	Product details
	Product details	Product details	Product details	Product details
	Product details	Product details	Product search	Product search
(52.51 to 194.99)	Notif. registered	Notif. registered	Notif. registered	Notif. registered
	Notif. registered	Notif. registered	Notif. registered	Notif. registered
	Notif. registered	Notif. registered	Rewards landing	Notif. registered
(195 and more)	Notif. registered	Notif. registered	Notif. registered	Notif. registered
	Notif. registered	Notif. registered	Notif. registered	Notif. registered
	Notif. registered	Notif. registered	Notif. registered	Notif. registered

Notes: Notif.- Notifications; features are listed by order of use. For example, rewards landing is the most frequent last feature, product details is the most frequent second-to-last feature, and rewards landing is the most frequent third-to-last feature prior to a purchase when the app is accessed within 6.57 hours and 1.3 miles of the store of purchase.

Table A23. Rewards Users' and Non-users' Purchases with and without Rewards

Variable	(1) Rewards purchases			(2) Non-Rewards purchases		
	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
Rewards users	0.74	1.61	42.39	1.29	1.22	66.11
Rewards non-users	0.50	1.08	30.24	0.69	2.15	37.08
Difference significance	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Note: "Rewards users" refers to shoppers who last used the rewards feature in the app before purchase.

Table A24. App Introduction and Product-related Purchase Mechanism

Variable	No. of unique products	No. of unique product categories	% of top 100 products	% of top 500 products
(App Adopters X Post App Introduction)	0.237*** (0.016)	0.501*** (0.037)	-0.006** (0.002)	-0.005* (0.003)
App Adopters	0.029 (0.022)	-0.007 (0.048)	-0.001 (0.002)	0.000 (0.002)
IMR	-2.265*** (0.064)	-4.463*** (0.159)	0.011*** (0.002)	0.016*** (0.004)
Intercept	7.98*** (0.17)	14.535*** (0.423)	0.087*** (0.003)	0.278*** (0.004)
Number of Observations	1,983,596	1,983,596	1,983,596	1,983,596
R squared	0.029	0.022	0.071	0.049

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included; *** p < 0.001, ** p < 0.01, * p < 0.10. NMV = Net Monetary Value. IMR = inverse Mills ratio from the selection model.

Table A25. Returns: Analysis by Original Channel of Purchase

Variable	Frequency of returns of offline purchases	Quantity of returns of offline purchases	Value of returns of offline purchases	Frequency of returns of online purchases	Quantity of returns of online purchases	Value of returns of online purchases
(App Adopters X Post App Introduction)	0.033*** (0.004)	0.045*** (0.008)	1.924*** (0.194)	0.004*** (0.001)	0.006*** (0.001)	0.289*** (0.06)
App Adopters	-0.007 (0.007)	-0.010 (0.011)	-0.021 (0.328)	0.001* (0.001)	0.002 (0.001)	0.087 (0.055)
IMR	-0.181*** (0.014)	-0.240*** (0.022)	-9.283*** (0.774)	-0.025*** (0.002)	-0.031*** (0.003)	-1.476*** (0.128)
Intercept	0.208*** (0.014)	0.271*** (0.022)	9.623*** (0.718)	0.026*** (0.002)	0.032*** (0.002)	1.434*** (0.125)
Number of Observations	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760
R squared	0.020	0.016	0.011	0.006	0.004	0.002

Notes: Robust standard errors clustered by zip code are in parentheses; month fixed effects are included *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. IMR = inverse Mills ratio from the selection model.

Figure A1. App Screenshots on iPhone

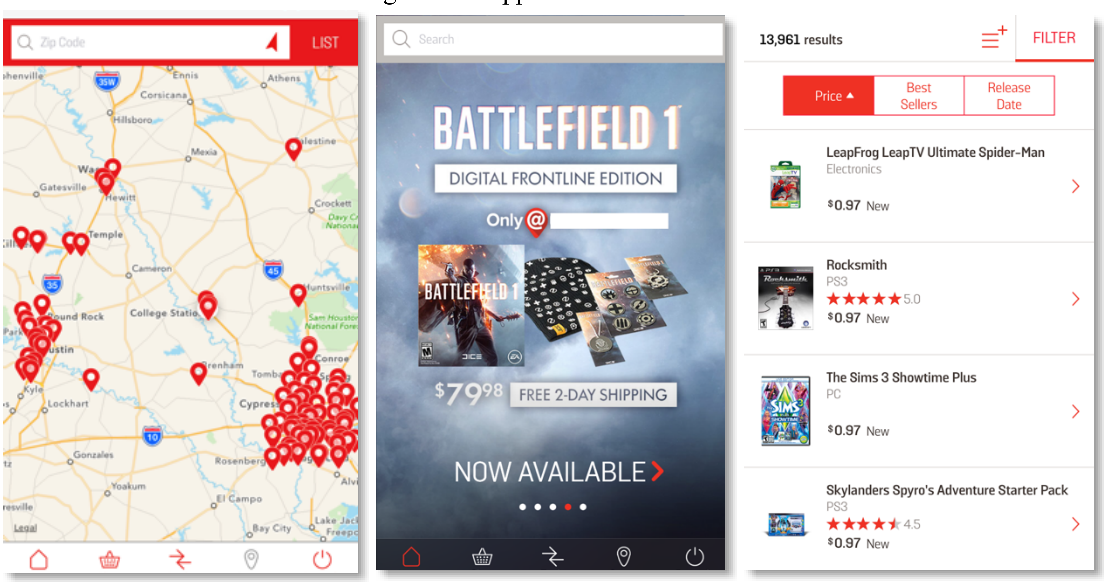


Figure A2. Pre-period Trends for Unmatched Sample

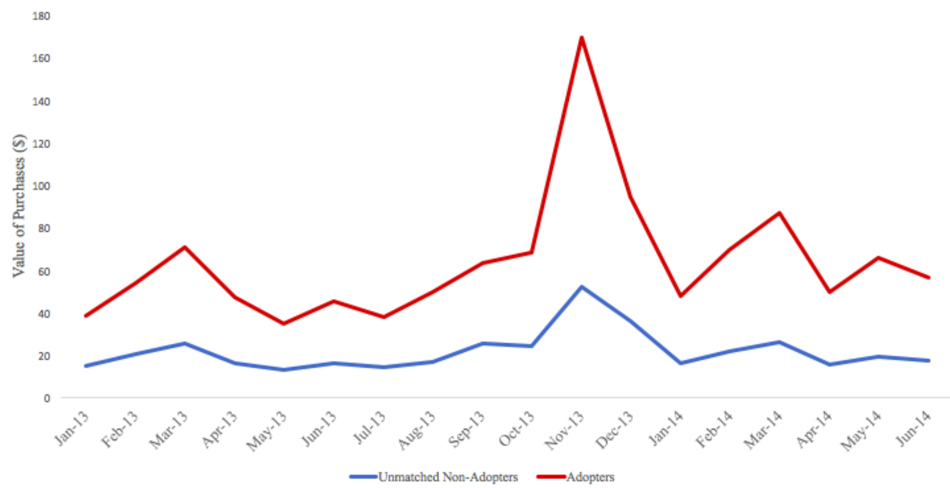


Figure A3. Period of Engagement:
Shopper Frequency for Number of Days App Accessed in the First Six Months

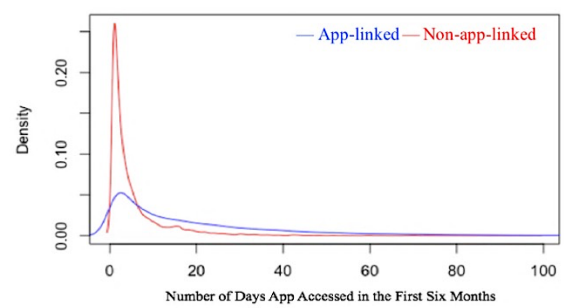


Figure A4. Period of Engagement: Shopper Frequency for
Range of Days (Last – First Day of App Access) in the First Six Months

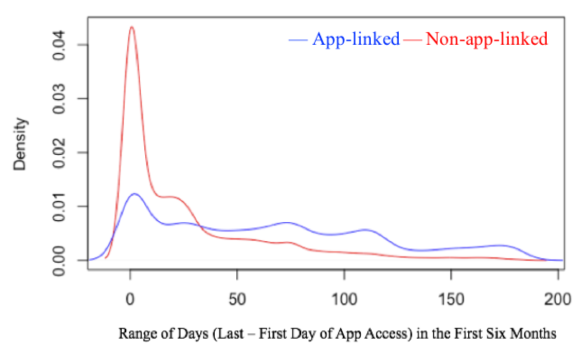


Figure A5. Number of App Sessions

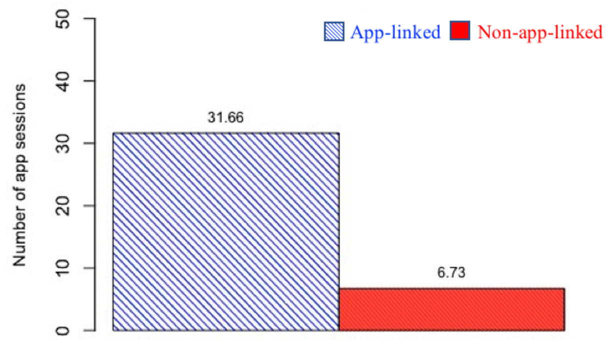


Figure A6. Dwell Time in App per Session

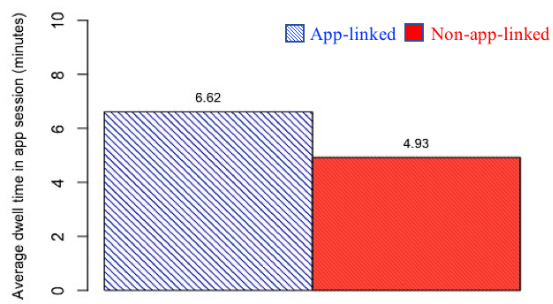


Figure A7. Number of Features Used within a Session

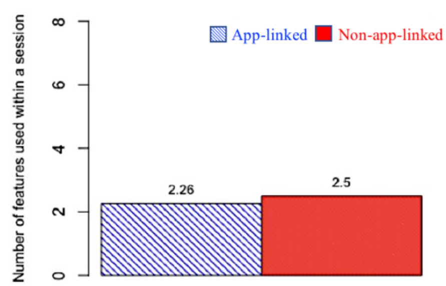
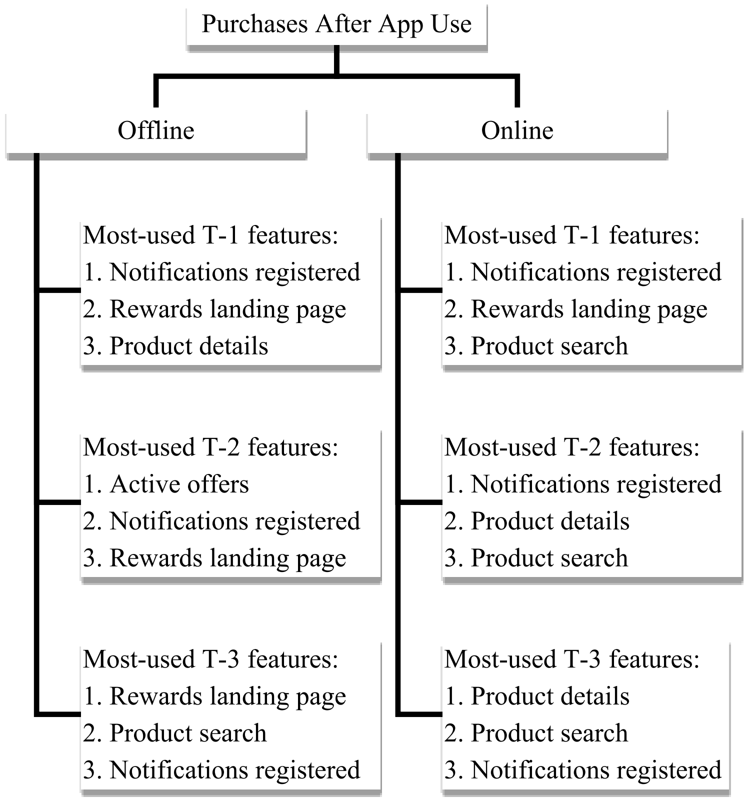


Figure A8. The Last Three Most Frequently Used Features before Purchase



Notes: T-1 refers to last feature used, T-2 refers to the second-to-last feature used, and T-3 refers to the third-to-last feature used in the app before purchase. Within each category, we show the highest frequency of use of feature. Thus, the most common last used features before an offline purchase are notifications, rewards landing, and product details in order of frequency of shoppers using these features.

APPENDIX B

APPENDIX I FOR CHAPTER IV

Causal Forest

Causal Trees: Overview. A causal tree is similar to a regression tree. The typical objective of a regression tree is to build accurate predictions of the outcome variable by recursively splitting the data into subgroups that differ the most on the outcome variable based on covariates. A regression tree has decision/internal/split nodes characterized by binary conditions on covariates and leaf or terminal nodes at the bottom of the tree. The regression tree algorithm continuously partitions the data, evaluating and re-evaluating at each node to determine (a) whether further splits would improve prediction, and (b) the covariate and the value of the covariate on which to split. The goodness-of-fit criterion used to evaluate the splitting decision at each node is the mean squared error (MSE) computed as the deviation of the observed outcome from the predicted outcome. The tree algorithm continues making further splits as long as the MSE decreases by more than a specified threshold.

The causal tree model adapts the regression tree algorithm in several ways to make it amenable for causal inference. First, it explicitly moves the goodness-of-fit-criterion to treatment effects rather than the MSE of the outcome measure. Second, it employs "honest" estimates, that is, the data on which the tree is built (splitting data) are separate from the data on which it is tested for prediction of heterogeneity (estimating data). Thus, the tree is honest if for a unit *i* in the training sample, it only uses the response Y_i to estimate the within-leaf treatment effect, or to decide where to place the splits, but not both (Athey and Imbens 2016; Athey et al. 2017). To avoid overfitting, we use cross-validation approaches in the tree-building stage.

Importantly, the goodness-of-fit criterion for causal trees is the difference between the estimated and the actual treatment effect at each node. While this criterion ensures that all the degrees of freedom are used well, it is challenging because we never observe the true treatment effect.

Causal Tree: Goodness-of-fit Criterion. Following Wager and Athey (2018), if we have *n* independent and identically distributed training examples labeled *i* = 1, ..., *n*, each of which consists of a feature vector $X_i \in [0, 1]^d$, a response $Y_i \in R$, and a treatment indicator $W_i \in [0, 1]$, the CATE at *x* is:

(2) $\tau(x) = \mathbb{E}\left[Y_i^1 - Y_i^0 | X_i = x\right]$

We assume unconfoundedness, i.e., conditional on X_i, the treatment W_i is independent of outcome Y_i. Because the true treatment effect is not observed, we cannot directly compute the goodness-of-fit criterion for creating splits in a tree. This goodness-of-fit criterion is as follows.

(3) $Q_{infeasible} = \mathbb{E}\left[((\tau_i(X_i) - \hat{\tau}_i(Xi))^2\right]$

Because $\tau_i(X_i)$ is not observed, we follow Athey and Imbens's (2016) approach to create a transformed outcome Y_i^* that represents the true treatment effect. Assume that the treatment indicator W_i is a random variable. Suppose there is a 50-50 probability for a unit *i* to be in the treated or the control group, an unbiased true treatment effect can be obtained for that unit by just using its outcomes Y in the following way. Let

(4) $Y_i^* = 2Y_i \, if \; W_i = 0$ and $Y_i^* = -2Y_i \, if \; W_i = 1$

It follows that:

(5) $\mathbb{E}[Y_i^*] = 2.(\frac{1}{2}\mathbb{E}[Y_i(1)] - \frac{1}{2}\mathbb{E}[Y_i(0)]) = \mathbb{E}[\tau_i]$

Therefore, we can compute the goodness-of-fit criterion for determining node splits in a causal tree using the expectation of the transformed outcome (see Athey and Imbens 2016 for details). Once we generate causal trees, we can compute the treatment effect within each leaf because it has a finite number of observations and standard asymptotics apply within a leaf. The differences in outcomes for the treated and control units within each leaf produces the treatment effect in that leaf.

Causal Forest Ensemble. In the final step, we create an ensemble of trees using ideas from model averaging and bagging. Specifically, we take predictions from thousands of trees and average over them (Guo et al. 2018). This step retains the unbiased, honest nature of tree-based estimates but reduces the variance. The forest averages over the estimates from B trees in the following manner.

(6) $\hat{\tau}(x) = B^{-1}\sum_{b=1}^{B}\hat{\tau}_b(x)$

Because monetary value of purchases is the key outcome variable of interest to the retailer, we estimate individual level treatment effect on value of purchases for each failure experiencer separately using the observed covariate data. These covariates include age, gender, loyalty program, mobile operating systems (OS) type, and the number of stores in the shopper's CBSA in addition to the three theoretically-driven moderators, namely, value of past purchases, recency of past purchases and online buying/digital experience. These individual attributes are important for identifying individual-level effects and for developing targeting approaches (e.g., Neumann et al. 2019). We use a random sample of two-thirds of our data as training data and the remaining one-third as test data for predicting CATE. We use half of the training data to maintain honest estimates and for cross-validation to avoid overfitting.

Table B1. Causal Forest Results: Summary of Individual Treatment Effect for Value of Purchases

(a) Pooled Sample

	N_{test}	Mean	SD	Min	Max
$\hat{\tau}$	96,071	-1.91	3.35	-13.41	27.59
$\hat{\tau} \mid \hat{\tau} < 0$	66,611	-3.42	2.62	-13.41	-0.00007
$\hat{\tau} \mid \hat{\tau} > 0$	29,460	1.53	2.04	0.00002	27.59

(b) April Sample

	N_{test}	Mean	SD	Min	Max
$\hat{\tau}$	8,942	-5.123	3.351	-17.101	4.537
$\hat{\tau} \mid \hat{\tau} < 0$	8,163	-5.722	2.845	-17.101	-0.001
$\hat{\tau} \mid \hat{\tau} > 0$	779	1.161	0.899	0.003	4.537

Note: $\hat{\tau}$ represents the estimated Conditional Average Treatment Effect (CATE) for each individual in the test data.

Table B2. Causal Forest: Post-hoc CATE Regression for Value of Purchases

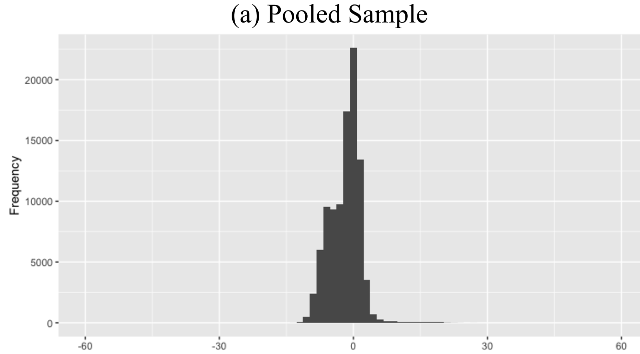
(a) Pooled Sample

	CATE (Standard Error)
Intercept	-1.313***
	(0.041)
Recency of purchases	-0.004***
	(0.000)
Past value of purchases	0.000***
	(0.000)
Past online purchase frequency	0.333***
	(0.006)
Age	-0.021***
	(0.001)
Gender	-0.046***
	(0.014)
Loyalty program	0.163***
	(0.017)
Number of stores in CBSA	0.002***
	(0.000)
Mobile operating system	-0.018
	(0.015)
R squared	0.61

(b) April Sample

	All CATE
Intercept	-2.218***
	(0.083)
Past value	-0.002***
	(0.000)
Online buyer or not	0.011
	(0.009)
Recency	-0.001***
	(0.0001)
Age	0.035***
	(0.002)
Gender	0.981***
	(0.039)
Loyalty program	0.052
	(0.040)
Number of Stores in CBSA	0.004***
	(0.0001)
Platform	-4.988***
	(0.036)
N	8,942
R squared	0.78

Figure B1. Causal Forest Results: Individual CATE
(a) Pooled Sample



(b) April Sample

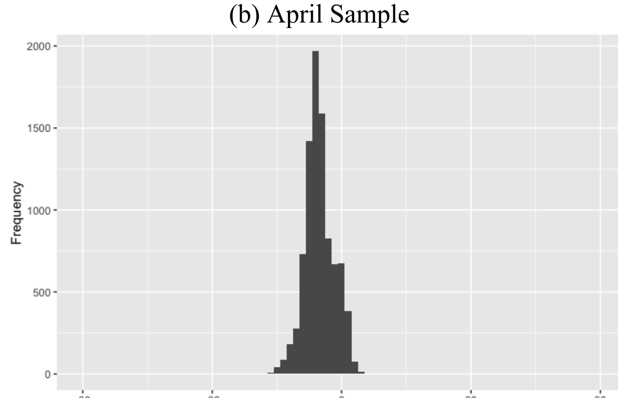
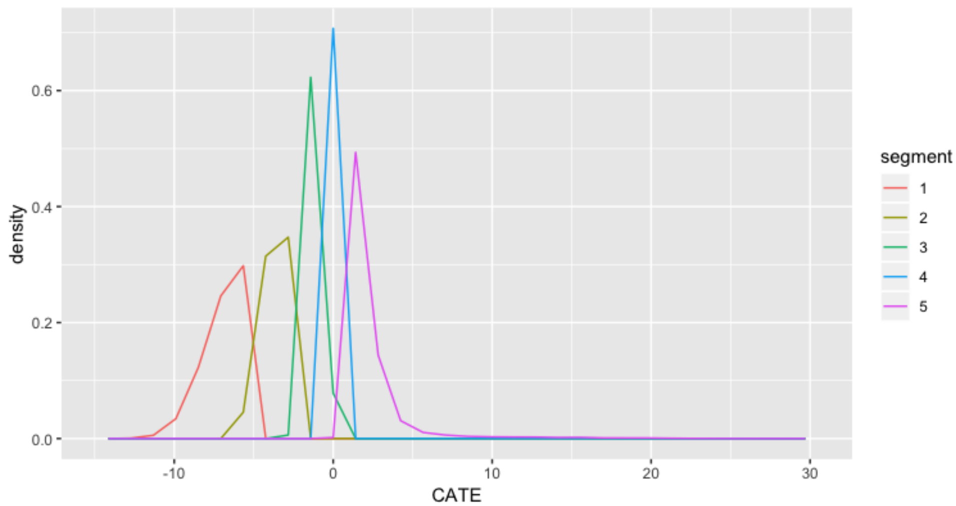
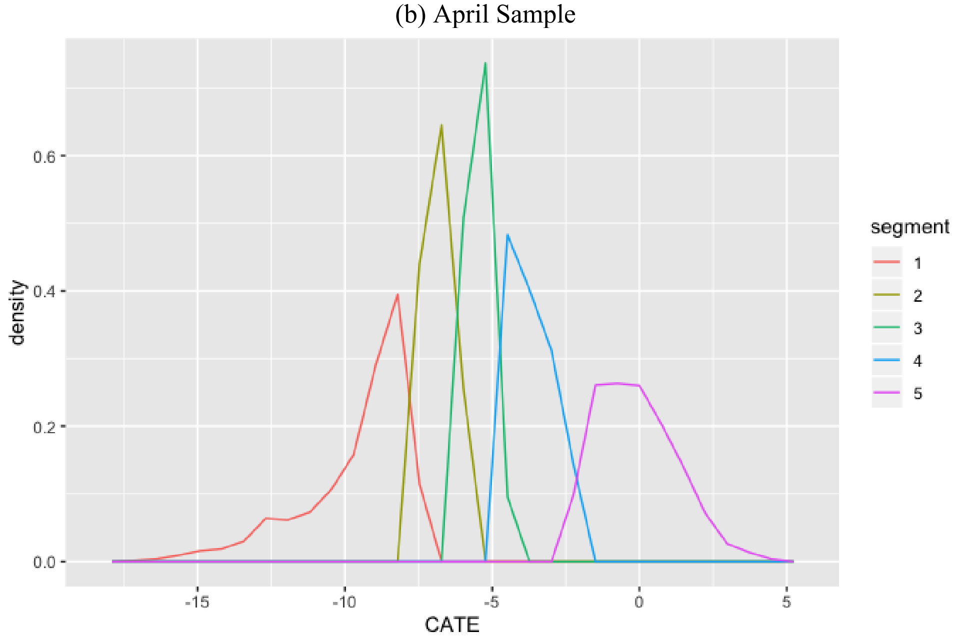


Figure B2. Causal Forest Results: Quintiles by CATE

(a) Pooled Sample



(b) April Sample



Note: Segment 1 represents shoppers most adversely affected by failure while Segment 5 represents those who are least adversely affected.

APPENDIX C

APPENDIX II FOR CHAPTER IV

Robustness Check for Table 16 (Main Treatment Effect) Results

In this section, we present the results for robustness checks for the main estimation in Table 16 relating to: (a) alternative models with Propensity Score Matching and using Poisson model (Tables C1-C2), (b) varying time periods (Table C3), (c) outliers (Table C4), (d) existing shoppers (Table C5), (e) alternative measures for prior use of digital channels (Table C6), (f) regression-discontinuity style analysis (Table C7), and (g) sample including multiple failure experiencers (Table C8).

Table C1. Robustness of Table 16 Results to Propensity Score Matching

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.041***	-0.100***	-3.117***
	(0.006)	(0.016)	(0.683)
Failure experiencers	0.044***	0.106***	4.279***
	(0.005)	(0.016)	(0.613)
Post shock	-0.028***	-0.117***	-2.839***
	(0.004)	(0.011)	(0.467)
Intercept	0.554***	1.003***	32.789***
	(0.006)	(0.015)	(0.556)
R squared	0.010	0.006	0.008

Notes: Number of observations is 537,772 for 134,443 treated shoppers with complete demographic data matched 1:1 with replacement out of a pool of 158,328 control shoppers with complete demographic data. Robust standard errors clustered by shoppers are in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. DID = Difference-in-Differences.

Table C2. Robustness of Table 16 Results to Poisson Specification of Count-based Outcomes

Variable	Frequency of purchases	Quantity of purchases
Failure experiencers x Post shock (DID)	-0.054*** (0.006)	-0.059*** (0.009)
Failure experiencers	0.005 (0.004)	0.004 (0.006)
Post shock	0.025*** (0.004)	-0.034*** (0.007)
Intercept	-0.016*** (0.004)	0.760*** (0.006)
Log pseudo-likelihood	-1,463,816	-2,985,536

Notes: Number of observations is 970,370. Robust standard errors in parentheses; *** p < 0.001, ** p < 0.01, * p < 0.05. DID = Difference-in-Differences.

Table C3. Robustness of Table 16 Results to 30-Day Time Period

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.092*** (0.01)	-0.175*** (0.032)	-8.051*** (1.025)
Failure experiencers	-0.044*** (0.007)	-0.126*** (0.023)	-1.805* (0.726)
Post shock	-0.099*** (0.007)	-0.429*** (0.022)	-7.473*** (0.709)
Intercept	2.039*** (0.006)	4.493*** (0.02)	113.289*** (0.627)
R squared	0.110	0.049	0.052

Notes: Number of observations is 970,370. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, ** p < 0.01, * p < 0.05. DID = Difference-in-Differences.

Table C4. Robustness of Table 16 Results to Outlier Spenders

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.051*** (0.006)	-0.108*** (0.014)	-3.033*** (0.458)
Failure experiencers	0.008* (0.004)	0.022* (0.010)	0.305 (0.325)
Post shock	0.077*** (0.004)	0.092*** (0.010)	18.09*** (0.317)
Intercept	0.897*** (0.003)	1.848*** (0.009)	42.074*** (0.28)
R squared	0.003	0.002	0.012

Notes: Number of observations is 934,638. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, ** p < 0.01, * p < 0.05. DID = Difference-in-Differences.

Table C5. Robustness of Table 16 Results to Existing Shoppers

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.053*** (0.006)	-0.116*** (0.020)	-3.81*** (0.724)
Failure experiencers	0.003 (0.004)	0.001 (0.014)	0.49 (0.513)
Post shock	0.028*** (0.004)	-0.063*** (0.014)	3.675*** (0.501)
Intercept	0.992*** (0.004)	2.143*** (0.012)	57.566*** (0.445)
R squared	0.002	0.001	0.005

Notes: Number of observations is 942,822. Robust standard errors clustered by shoppers are in parentheses; *** $p < 0.001$. DID = Difference-in-Differences.

Table C6. Robustness of Table 16 Results to Alternative Measures of Digital Channel Use based on App Usage Before Failure

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.474*** (0.007)	-1.107*** (0.022)	-35.032*** (0.817)
DID x Past value of purchases	0.000*** (0.000)	0.001*** (0.000)	0.032*** (0.000)
DID x Recency of purchases	-0.002*** (0.000)	-0.005*** (0.000)	-0.175*** (0.005)
DID x Past app use frequency	-0.019*** (0.001)	-0.052*** (0.002)	-1.169*** (0.064)
Past value of purchases	0.000*** (0.000)	0.000*** (0.000)	0.011*** (0.000)
Recency	0.003*** (0.000)	0.006*** (0.000)	0.2*** (0.002)
Past online purchase frequency	0.064*** (0.000)	0.146*** (0.001)	3.851*** (0.032)
Failure experiencers	0.064*** (0.004)	0.171*** (0.013)	5.106*** (0.487)
Post shock	0.022*** (0.004)	-0.082*** (0.013)	2.711*** (0.475)
Intercept	0.791*** (0.004)	1.466*** (0.013)	40.48*** (0.466)
R squared	0.1480	0.1149	0.0797

Notes: Number of observations is 964,916. Robust standard errors clustered by shoppers are in parentheses; Each moderator interacts with the difference-in-differences (DID) term failure experiencers x post shock; *** $p < 0.001$. The observations include those of shoppers with at least one purchase in the past for computing recency.

Table C7. Robustness of Table 16 Results to Regression Discontinuity Style Analysis

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.069***	-0.183***	-2.606
	(0.008)	(0.027)	(1.479)
Failure experiencers	0.061***	0.137***	2.740
	(0.008)	(0.031)	(1.432)
Post shock	0.065***	0.019	5.115***
	(0.006)	(0.024)	(1.382)
Intercept	0.911***	1.974***	53.976***
	(0.007)	(0.027)	(1.182)
R squared	0.001	0.001	0.003

Notes: Number of observations is 345,708. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001. DID = Difference-in-Differences.

Table C8. Robustness of Table 16 Results for Shoppers with Multiple Failures

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.049***	-0.086***	-3.603***
	(0.004)	(0.012)	(0.552)
Failure experiencers	0.047***	0.112***	2.960***
	(0.004)	(0.013)	(0.513)
Post shock	0.025***	-0.075***	2.997***
	(0.003)	(0.009)	(0.454)
Intercept	0.950***	2.048***	56.110***
	(0.004)	(0.014)	(0.456)
R squared	0.002	0.002	0.005

Notes: Number of observations is 1,190,056. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001. DID = Difference-in-Differences.

APPENDIX D

APPENDIX III FOR CHAPTER IV

Robustness Check for Table 17 (By Channel) Results

In this section, we present the results for robustness checks for the cross-channel estimation in Table 17 relating to (a) alternative models with Propensity Score Matching and using Poisson model (Tables D1-D2), (b) varying time periods (Table D3), (c) outliers (Table D4), (d) existing shoppers (Table D5), (e) alternative measures for prior use of digital channels (Table D6), (f) regression-discontinuity style analysis (Table D7), and (g) sample including multiple failure experiencers (Table D8).

Table D1. Robustness of Table 17 Results to Propensity Score Matching

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.039*** (0.006)	-0.098*** (0.016)	-2.941*** (0.668)	-0.001 (0.001)	-0.003 (0.002)	-0.176 (0.136)
Failure experiencers	0.041*** (0.005)	0.100*** (0.016)	3.838*** (0.599)	0.003** (0.001)	0.006** (0.002)	0.441*** (0.116)
Post shock	-0.019*** (0.004)	-0.101*** (0.011)	-2.121*** (0.459)	-0.009*** (0.001)	-0.016*** (0.002)	-0.718*** (0.080)
Intercept	0.522*** (0.006)	0.959*** (0.015)	31.028*** (0.546)	0.032*** (0.001)	0.043*** (0.002)	1.762*** (0.101)
R squared	0.009	0.005	0.008	0.002	0.002	0.001

Notes: Number of observations is 537,772 for 134,443 treated shoppers with complete demographic data matched 1:1 with replacement out of a pool of 158,328 control shoppers with complete demographic data. Robust standard errors clustered by shoppers are in parentheses; *** $p < 0.001$, ** $p < 0.01$. DID = Difference-in-Differences.

Table D2. Robustness of Table 17 Results to Poisson Specification for Count Outcomes

	(1) Offline		(2) Online	
Variable	Frequency of purchases	Quantity of purchases	Frequency of purchases	Quantity of purchases
Failure experiencers x Post shock (DID)	-0.055***	-0.060***	-0.044	-0.014
	(0.006)	(0.009)	(0.029)	(0.051)
Failure experiencers	0.0045	0.004	0.003	0.016
	(0.0043)	(0.006)	(0.019)	(0.029)
Post shock	0.0302***	-0.032***	-0.098***	-0.116**
	(0.0042)	(0.007)	(0.020)	(0.035)
Intercept	-0.0646***	0.722***	-3.050***	-2.5420***
	(0.0039)	(0.006)	(0.016)	(0.0273)
Log pseudo-likelihood	-1,432,743	-2,941,566	-162,877	-254,328

Notes: Number of observations is 970,370. Robust standard errors in parentheses; *** p < 0.001, ** p < 0.01. DID = Difference-in-Differences.

Table D3. Robustness of Table 17 Results to 30-Day Time Period

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.087***	-0.167***	-7.652***	-0.004*	-0.009	-0.399**
	(0.009)	(0.031)	(1.008)	(0.002)	(0.006)	(0.142)
Failure experiencers	-0.042***	-0.126***	-1.942**	-0.002	0.000	0.136
	(0.007)	(0.022)	(0.714)	(0.001)	(0.004)	(0.101)
Post shock	-0.098***	-0.429***	-7.494***	-0.001	0.001	0.021
	(0.006)	(0.022)	(0.697)	(0.001)	(0.004)	(0.098)
Intercept	1.943***	4.313***	108.151***	0.096***	0.18***	5.138***
	(0.006)	(0.019)	(0.617)	(0.001)	(0.004)	(0.087)
R squared	0.108	0.049	0.050	0.007	0.002	0.004

Notes: Number of observations is 970,370. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, ** p < 0.01, * p < 0.05. DID = Difference-in-Differences.

Table D4. Robustness of Table 17 Results to Outlier Spenders

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.049*** (0.005)	-0.107*** (0.014)	-2.960*** (0.448)	-0.002 (0.001)	-0.001 (0.002)	-0.073 (0.085)
Failure experiencers	0.008* (0.004)	0.021* (0.010)	0.232 (0.317)	0.000 (0.001)	0.001 (0.002)	0.073 (0.060)
Post shock	0.078*** (0.004)	0.094*** (0.010)	17.691*** (0.310)	-0.002* (0.001)	-0.003 (0.001)	0.399*** (0.059)
Intercept	0.854*** (0.003)	1.781*** (0.009)	39.821*** (0.273)	0.043*** (0.001)	0.067*** (0.001)	2.254*** (0.052)
R squared	0.0024	0.0017	0.0118	0.0011	0.0010	0.0008

Notes: Number of observations is 934,638. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, * p < 0.05. DID = Difference-in-Differences.

Table D5. Robustness of Table 17 Results to Existing Shoppers

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of Purchases
Failure experiencers x Post shock (DID)	-0.051*** (0.006)	-0.115*** (0.020)	-3.628*** (0.713)	-0.001 (0.001)	-0.001 (0.003)	-0.182 (0.108)
Failure experiencers	0.003 (0.004)	0.000 (0.014)	0.306 (0.505)	0.000 (0.001)	0.001 (0.002)	0.184* (0.076)
Post shock	0.032*** (0.004)	-0.055*** (0.014)	3.743*** (0.493)	-0.004*** (0.001)	-0.007*** (0.002)	-0.067 (0.075)
Intercept	0.945*** (0.004)	2.064*** (0.012)	54.784*** (0.438)	0.048*** (0.001)	0.079*** (0.002)	2.783*** (0.066)
R squared	0.0019	0.0012	0.0045	0.0012	0.0007	0.0007

Notes: Number of observations is 942,822. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, * p < 0.05. DID = Difference-in-Differences.

Table D6. Robustness of Table 17 Results to Alternative Measure of Digital Channel Use based on App Usage Before Failure

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of Purchases
Failure experiencers x Post shock (DID)	-0.455*** (0.007)	-1.027*** (0.022)	-33.232*** (0.805)	-0.019*** (0.001)	-0.08*** (0.003)	-1.799*** (0.127)
DID x Past value of purchases	0.000*** (0.000)	0.001*** (0.000)	0.031*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.002*** (0.000)
DID x Recency of purchases	-0.002*** (0.000)	-0.005*** (0.000)	-0.164*** (0.005)	0.000*** (0.000)	0.000*** (0.000)	-0.01*** (0.001)
DID x Past app use frequency	-0.018*** (0.001)	-0.051*** (0.002)	-1.128*** (0.064)	-0.001*** (0.000)	0.000 (0.000)	-0.041*** (0.010)
Past value of purchases	0.000*** (0.000)	0.000*** (0.000)	0.011*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
Recency	0.003*** (0.000)	0.006*** (0.000)	0.193*** (0.002)	0.000*** (0.000)	0.000*** (0.000)	0.007*** (0.000)
Past online purchase frequency	0.060*** (0.000)	0.139*** (0.001)	3.602*** (0.032)	0.004*** (0.000)	0.007*** (0.000)	0.248*** (0.005)
Failure experiencers	0.061*** (0.004)	0.164*** (0.013)	4.695*** (0.48)	0.003*** (0.001)	0.007*** (0.002)	0.411*** (0.075)
Post shock	0.026*** (0.004)	-0.075*** (0.013)	2.769*** (0.468)	-0.004*** (0.001)	-0.007** (0.002)	-0.058 (0.074)
Intercept	0.758*** (0.004)	1.417*** (0.013)	38.674*** (0.46)	0.033*** (0.001)	0.049*** (0.002)	1.806*** (0.072)
R squared	0.1423	0.1115	0.0771	0.0117	0.0090	0.0058

Notes: Number of observations is 964,916. Robust standard errors clustered by shoppers are in parentheses; *** $p < 0.001$. DID = Difference-in-Differences. The observations include those of shoppers with at least one purchase in the past for computing recency.

Table D7. Robustness of Table 17 Results to Regression Discontinuity Style Analysis

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.071***	-0.189***	-2.587*	0.002	0.006	-0.019
	(0.008)	(0.027)	(1.465)	(0.002)	(0.004)	(0.181)
Failure experiencers	0.06***	0.135***	2.478	0.001	0.002	0.262
	(0.007)	(0.031)	(1.423)	(0.001)	(0.003)	(0.148)
Post shock	0.071***	0.031	5.110***	-0.006***	-0.012***	0.006
	(0.006)	(0.023)	(1.374)	(0.001)	(0.003)	(0.134)
Intercept	0.866***	1.903***	51.516***	0.044***	0.071***	2.461***
	(0.007)	(0.027)	(1.175)	(0.001)	(0.003)	(0.109)
R squared	0.001	0.0007	0.0026	0.0006	0.0006	0.0002

Notes: Number of observations is 345,708. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, * p < 0.10. DID = Difference-in-Differences.

Table D8. Robustness of Table 17 Results to Multiple Failures

	(1) Offline			(2) Online		
Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x Post shock (DID)	-0.046***	-0.083***	-3.310***	-0.003**	-0.003	-0.292**
	(0.004)	(0.012)	(0.544)	(0.001)	(0.002)	(0.089)
Failure experiencers	0.045***	0.109***	2.669***	0.002**	0.004*	0.291***
	(0.004)	(0.013)	(0.507)	(0.001)	(0.002)	(0.069)
Post shock	0.029***	-0.069***	3.049***	-0.004***	-0.007***	-0.051
	(0.003)	(0.009)	(0.449)	(0.001)	(0.002)	(0.064)
Intercept	0.904***	1.972***	53.41***	0.046***	0.076***	2.700***
	(0.004)	(0.013)	(0.450)	(0.001)	(0.002)	(0.06)
R squared	0.0017	0.0015	0.0046	0.001	0.0007	0.0007

Notes: Number of observations is 1,190,056. Robust standard errors clustered by shoppers are in parentheses; *** p < 0.001, ** p < 0.01, * p < 0.05. DID = Difference-in-Differences.

APPENDIX E

APPENDIX IV FOR CHAPTER IV

Table E1. Persistence of Effects of App Failure on Average Value of Purchases

Variable	Value of purchases
Treat x Week -4	-0.09
	(0.28)
Treat x Week -3	-0.58*
	(0.25)
Treat x Week -2	-0.45
	(0.24)
Treat x Week -1	-0.37
	(0.25)
Treat x Week 1	-0.82*
	(0.37)
Treat x Week 2	-1.7**
	(0.56)
Treat x Week 3	-0.58[1]
	(0.31)
Treat x Week 4	-0.53[1]
	(0.31)
Treat x Week 5	-0.64*
	(0.29)
Intercept	13.19***
	(0.09)

Notes: Robust standard errors clustered by shoppers are in parentheses; week and individual fixed effects are included. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, [1] $p < 0.1$. N = 1,366,890. DID = Difference-in-Differences.

Table E2. Stacked Online and Offline Purchases for Cross-Channel Effects

Variable	Frequency of purchases	Quantity of purchases	Value of purchases
Failure experiencers x post shock (DID)	-0.001	-0.003	-0.093
	(0.002)	(0.003)	(0.154)
DID x Channel	-0.021**	-0.052**	-1.994**
	(0.008)	(0.019)	(0.674)
Failure experiencers	-0.003*	-0.005*	-0.167**
	(0.001)	(0.002)	(0.064)
Post shock	0.009***	0.015***	1.672***
	(0.001)	(0.002)	(0.113)
Channel	0.678***	1.416***	27.755***
	(0.005)	(0.012)	(0.217)
Failure experiencers x Channel	-0.015*	-0.02	-0.36
	(0.006)	(0.017)	(0.299)
Post shock x Channel	0.161***	0.206***	23.602***
	(0.006)	(0.014)	(0.493)
Intercept	0.036***	0.054***	1.5***
	(0.001)	(0.002)	(0.048)
R squared	0.1398	0.0949	0.0911

Notes: Robust standard errors clustered by shoppers are in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. N = 546,756. DID = Difference-in-Differences. Channel is 1 for store purchases and 0 for online purchases.