Essays on Omnichannel Retailing: Customer Channel Adoption and Pricing Implications

Haonan Zhang

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Shirsho Biswas, Chair

Hema Yoganarasimhan, Chair

Simha Mummalaneni

Oliver J. Rutz

Program Authorized to Offer Degree:
Business Administration

University of Washington

Abstract

Essays on Omnichannel Retailing: Customer Channel Adoption and Pricing Implications

Haonan Zhang

Co-Chairs of the Supervisory Committee:
Assistant Professor Shirsho Biswas
Marketing and International Business

Professor Hema Yoganarasimhan
Marketing and International Business

The rapid growth of digital shopping channels has prompted many traditional retailers to invest in e-commerce websites and mobile apps. While prior literature shows that multichannel customers are more valuable, it has paid less attention to how the motive for adopting a new channel shapes post-adoption behavior. In this dissertation, I study consumer behavior and firm strategy in multichannel retailing. Chapter 2 examines how different online-channel adoption pathways relate to post-adoption customer behavior and customer value. Chapter 3 studies how the adoption of a retailer's online channel affects consumers' subsequent responses to price changes during offline purchase occasions, that is, offline price sensitivity. Together, these chapters highlight the importance of incorporating online-channel adoption behavior and heterogeneous adoption pathways into omnichannel retailers' decisions on channel management, customer value forecasting, and targeting, with the goal of improving long-term customer value creation.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

As I look back on the past five years, I feel very fortunate to have been surrounded by people who supported, encouraged, and guided me throughout this journey. I am deeply grateful to my advisors, committee members, faculty mentors, peers, colleagues, friends, and family. This Ph.D. would not have been possible without their help.

First, I would like to thank my dissertation co-chairs, Shirsho Biswas and Hema Yoganarasimhan. Both of you have supported me in every part of my academic, professional, and personal growth. You took a chance on me, guided me through many difficult moments, and helped me grow as a researcher and as a person.

Hema, I first met you in the dynamic structural model seminar. That class opened my eyes to a new way of thinking about research. Since then, your way of approaching problems, your sharp thinking, your direct communication, and your high standards have shaped the way I think about research and professional work. Thank you for believing in me, challenging me, and showing me what it means to be both an outstanding scholar and an exceptional mentor. I am especially grateful for your support not only in my academic work, but also in my industry work and long-term career development.

Shirsho, I first worked with you on one of my earliest research projects, and your guidance has been a strong source of support ever since. I learned a great deal from the way you think through research questions, tackle challenges, and respond quickly and carefully to problems. Your feedback and support helped me move the paper forward and gave me confidence during many difficult stages of the Ph.D. journey. Thank you for your patience, guidance, and constant support.

I would also like to thank the other members of my supervisory committee, Simha Mummalaneni, Oliver J. Rutz, and Quan Wen. Simha, thank you for your constructive feedback on my research and for your guidance during the job market process. Oliver, thank you for helping me deepen my understanding of empirical marketing and for giving me thoughtful advice on my research. Quan, thank you for your microeconomics class, which helped me build a stronger foundation in economic thinking and has been very useful for my research.

I am also grateful to the Wiley Fellowship for the generous financial support and encouragement to pursue quantitative marketing research. I would like to thank all the Marketing faculty members at Foster for their support throughout my Ph.D. journey and especially during the job market. I also want to thank Beau Kirkeby for his continuous help and support at every stage of the program.

I am very thankful to all my peers and colleagues in the Marketing Department. I learned so much from being around such smart and generous people. Thank you for sharing your ideas, giving feedback on my work, and helping me prepare for job talks and interviews. In addition, I also want to thank Xiaoxuan Hou for her continuous support throughout the five-year Ph.D. journey. I am also grateful to Mengjie (Magie) Cheng, Ye Liu, Ziqian Liu, Zhiwei (Simon) Luo, Omid Rafieian, Taotao Ye, Jingwen Zhang, and Zihan Zhang for their support during the job market process.

Finally, I would like to thank my dear parents, Hui Zhang and Guiqing Du. Thank you for your unconditional love, support in every possible way, and constant encouragement. Thank you for staying up late with me and rehearsing my job talks more than fifty times. Your support gave me the strength to keep going, especially during the most difficult moments. I could not have completed this journey without you.

DEDICATION

To my beloved parents, Hui Zhang and Guiqing Du,
for their unconditional love and support.

Chapter 1
INTRODUCTION

The rapid growth of digital shopping channels has prompted many traditional retailers to invest in e-commerce websites and mobile apps. This shift has transformed consumer behavior and raised important questions about how online adoption affects customers' subsequent shopping decisions across channels. While prior literature shows that multichannel customers are more valuable than single-channel customers, it has paid less attention to two important questions. First, how does the motive for adopting a new channel shape post-adoption behavior? Second, given the significant differences between online and offline channels, such as search costs, do consumers who adopt a retailer's online channel become more price sensitive in their subsequent *offline* purchases with that retailer?

In this thesis, I study these questions in the context of multichannel retailing. Chapter 2 examines heterogeneity in online adoption pathways and post-adoption behaviors. Chapter 3 studies one important behavioral consequence of online adoption: changes in offline price sensitivity. Together, these chapters show that online adoption is associated not only with changes in customer value and channel usage, but also with changes in how consumers respond to offline prices.

- **Chapter 2 – Channel Adoption Pathways and Post-Adoption Behavior.** Using transaction-level data from a major Brazilian pet supplies retailer, I study offline-only consumers who adopt online shopping via four distinct pathways: organic adoption, the COVID-19 pandemic, Black Friday promotions, and a loyalty program. I examine how these different pathways are associated with differences in post-adoption spend, profitability, and channel usage using consumer-level panel data and difference-in-differences estimates. I find that all adopters increase spending relative to offline-only consumers, but their post-adoption behaviors differ systematically by adoption motive. Promotion-driven adopters engage in forward buying and exhibit lower subsequent profitability, whereas COVID-19 adopters display stronger offline persistence, consistent with consumer inertia and habit theory. These findings have important managerial implications: firms should design promotions that discourage stockpiling, reinforce habits among customers pushed online by external shocks, and explicitly account for heterogeneity in channel adoption motives when forecasting customer lifetime value and assessing the breakeven and ROI of promotions designed to induce new-channel adoption. This chapter is based on the paper *Channel Adoption Pathways and Post-Adoption Behaviors* (Biswas et al., 2025*a*).

- **Chapter 3 – How Does Online Shopping Affect Offline Price Sensitivity.** Using transaction-level data from the same Brazilian pet supplies retailer, I study how the adoption of a retailer's online channel affects consumers' price sensitivity during subsequent offline shopping occasions. I compare "adopters" – customers who began shopping online after a period of offline-only purchasing – with "non-adopters" who remained offline-only. I estimate a discrete choice logit model with individual-level heterogeneity, based on an algorithm that can handle both high-dimensional fixed effects and price endogeneity. I then apply a staggered difference-in-differences approach to the estimated price elasticities and obtain the Average Treatment Effect on the Treated (ATT). I find that offline price sensitivity increases significantly after online adoption in three out of four product categories, particularly in items with low switching costs, such as pet hygiene. These results underscore the importance of recognizing cross-channel effects in consumer behavior and contribute to the literature on pricing and multichannel retailing by identifying online adoption as a key driver of offline price sensitivity. This chapter is based on the paper *How Does Online Adoption Affect Offline Price Sensitivity* (Biswas et al., 2025*b*).

While both chapters use the same empirical setting and data source, the analysis samples differ according to the purpose of each chapter. Therefore, I present the setting and data separately in each chapter to keep the chapters self-contained. Some information may appear redundant, but this structure allows each chapter to be read independently, with its contributions and implications discussed separately.

Keywords: Multichannel Retailing, Online Shopping, Customer Value, Channel Adoption, Digital Platforms, Price Elasticity, Pricing, High-Dimensional Fixed Effects

Chapter 2

CHANNEL ADOPTION PATHWAYS AND POST-ADOPTION BEHAVIOR

2.1 Introduction

Background and Motivation

Online shopping has grown rapidly over the last two decades with the proliferation of computers and smartphones. In response, many grocery and retail chains (e.g., Walmart, Target, Costco), historically brick-and-mortar, now also sell through websites and mobile apps. For traditional retailers, a central question is not only whether online adoption increases consumer value, but also whether consumers who adopt online shopping through different pathways are equally valuable afterward (Goyal, 2023; Bloomberg, 2020).

Prior research has found that multichannel consumers (those who buy using multiple channels of a retailer) tend to spend more than single-channel consumers; see §2.2. However, this work largely ignores *why* consumers adopt new channels, implicitly assuming that the benefits from multichannel adoption are the same regardless of the reasons for adoption of new channels. This assumption is unlikely to hold in practice. Some customers adopt online shopping organically because they value the convenience of an additional channel. Others adopt because an external shock pushes them online, or because a firm-driven promotion requires online engagement. These distinct adoption pathways can attract different customers, induce different behaviors around adoption, and generate different post-adoption trajectories.

Understanding this heterogeneity is managerially important. Firms often evaluate the ROI of online adoption campaigns by extrapolating from the behavior of existing multichannel customers or organic adopters. Such extrapolation is valid only if customers induced to adopt by such campaigns behave like organic adopters after adoption. If, instead, promotion-induced or macro shock-induced adopters differ in spend, channel usage, or profitability, then firms could make errors in forecasting customer value, overpay to induce online adoption, and misjudge the ROI of campaigns designed to move customers online. These considerations motivate our research questions.

Research Questions and Theoretical Predictions

We study two research questions. First, how does adopting a retailer's digital shopping channels affect the purchase behavior of customers who were previously *offline-only* shoppers? Second, conditional on adoption, how does post-adoption behavior differ across organic, macro-shock-driven, and promotion-driven adoption pathways? We focus on three managerially salient outcomes: total spend, profitability, and the share of spending allocated to the offline channel.

Building on the multichannel literature and consumer behavior frameworks (discussed in detail in §2.2 and §2.3), we hypothesize that customers who adopt a firm's online channels and become multichannel shoppers will, on average, increase their total spending relative to *offline-only customers*. Beyond this average effect, we anticipate substantial heterogeneity across adoption pathways. Consumer inertia and habit theory (Verplanken and Wood, 2006; Wood and Neal, 2009) suggest that when adoption is induced by external shocks or incentives rather than intrinsic preference, consumers are more likely to revert toward prior routines once the stimulus subsides. Thus, relative to externally driven adopters, organically adopting customers should exhibit a higher post-adoption share of spending on online channels and faster growth in that share. We further anticipate that such differences will be more pronounced for older, habit-driven customers.

For adoption spurred by time-limited promotions, theories of forward buying and demand distortion (Neslin et al., 1985; Hendel and Nevo, 2006; Van Heerde and Neslin, 2017) imply that promotion-driven adopters may temporarily inflate their spending during the promotional period by stocking up, but then reduce spending afterwards. This mechanism can dampen or even offset the long-run spending uplift typically associated with multichannel adoption. Finally, because profit margins generally differ across online and offline channels, heterogeneity in the post-adoption mix of spending across channels is likely to translate into systematic differences in profitability across adopter types, even when their total spend is similar.

In summary, our theoretical framework predicts both an average multichannel effect and systematic pathway heterogeneity. Adopters should spend more and become more profitable than *offline-only customers*, but the magnitude and persistence of these gains should depend on why customers adopted online shopping. Macro-shock-driven adopters may exhibit stronger offline persistence due to inertia, while promotion-driven adopters may show lower post-adoption spend because promotions pull demand forward. These behavioral differences should, in turn, affect profitability through both total spend and channel mix.

Empirical Setting and Methodology

To test these predictions, we use data from a large pet supplies retailer in Brazil with an extensive network of brick-and-mortar stores and an online presence (website and mobile app). Our data comprise all customer purchases across all channels from 2019 to 2024. During this period, there were three key events: a macro-environmental shock (COVID-19) and two firm-driven promotions designed to incentivize online shopping. This allows us to observe a variety of adoption pathways.

We construct four groups of adopters of online shopping who previously only shopped at the firm's physical stores: (1) *covid adopters*, who started shopping online during the onset of COVID-19; (2) *Black Friday adopters*, who adopted online shopping to take advantage of deep discounts available online during Black Friday; (3) *Loyalty Program adopters*, who adopted online shopping during the launch of a loyalty program that had to be activated on the retailer's mobile app; and (4) *organic adopters*, who started shopping online without any identifiable external stimulus. In addition, we consider a control group of non-adopters or *offline-only customers* (consumers who did not adopt online shopping during our observation period).

Using these data, we construct consumer-level monthly panel data to quantify revenue, profitability, and the share of spend in the offline and online channels, and use this panel to answer our research questions. Specifically, we measure the Average Treatment Effect on the Treated (ATT) of online adoption on post-adoption behavior for different adoption pathways using a difference-in-differences framework with individual and time fixed effects. We interpret these estimates as descriptive ATTs with appropriate controls. This estimand is managerially relevant because each pathway reflects a different form of customer self-selection into online adoption. Firms typically cannot force adoption; they can only encourage it through campaigns or observe adoption triggered by external shocks. The relevant managerial object is therefore the post-adoption value of customers who actually adopt through each pathway, rather than a causal ATE for a randomly assigned adoption intervention.

Main Findings, Managerial Implications, and Contributions

Table 2.1 summarizes the main post-adoption differences between each externally induced adopter group and *organic adopters*. The table provides a qualitative summary of the key results and mechanisms; exact DID estimates, effect-size calculations, and statistical tests are reported in §2.6.2.

We find three main patterns. First, all adopters spend more and generate higher profits than comparable *offline-only customers*, confirming the average value of multichannel adoption. However, this pooled adopter comparison masks substantial pathway heterogeneity. Second, relative to *organic adopters*, *covid adopters* exhibit similar post-adoption spend but higher profitability. This profitability advantage is consistent with their greater reliance on the offline channel, where margins are higher in our setting, and with slower movement toward online shopping among older, more habit-driven customers.

Third, promotion-induced adopters behave differently from both *covid adopters* and *organic adopters*. *Black Friday adopters* and *Loyalty Program adopters* spend less than *organic adopters* after adoption and are also less profitable. Event-study evidence shows that these

Table 2.1. Summary of Main Findings by Adoption Pathway

Adoption pathway	*Covid adopters*	*Black Friday adopters*	*Loyalty Program adopters*
Post-adoption spend	Similar to *organic adopters*	Lower than *organic adopters*	Lower than *organic adopters*
Buying behavior around adoption	No evidence of forward buying	Forward buying during the promotion, followed by lower post-adoption spending	Forward buying around adoption, followed by lower post-adoption spending
Post-adoption channel mix	Higher offline share and slower movement toward online shopping	Higher offline share initially, but faster movement toward online shopping over time	No meaningful difference in offline share relative to *organic adopters*
Post-adoption profitability	Higher than *organic adopters*, despite similar spend	Lower than *organic adopters*	Lower than *organic adopters*
Main behavioral mechanism	Consumer inertia and habit persistence, especially among older customers; higher offline margins raise profitability	Promotion-induced forward buying reduces later spend; higher offline share does not offset the spending decline	Promotion-induced forward buying reduces later spend; recurring app-based activation may reinforce online use
Managerial implication	Shock-induced adopters can be valuable, but may need targeted reinforcement to develop online habits	Discount-induced adoption can overstate long-run customer value if managers ignore post-promotion demand pull-forward	Loyalty-program adoption can create online engagement, but firms should not assume these adopters behave like *organic adopters*

Notes: This table summarizes the main directional findings and mechanisms from the event-adopter versus organic-adopter analyses. The exact DID estimates, effect sizes, and statistical tests are reported in §2.6.2.

customers spend more around the adoption period but less afterward, consistent with forward buying and post-promotion demand distortion. The two promotion pathways also differ in channel usage. *Black Friday adopters* initially retain a higher offline share than *organic adopters* but move online faster over time, consistent with their younger profile. In contrast, *Loyalty Program adopters* show no meaningful difference in offline spending share, likely because monthly app-based activation reinforces online engagement.

Taken together, these findings show that the *reason* for online adoption matters. Macro-shock adopters can be more profitable than *organic adopters* when they continue using higher-margin offline channels, whereas promotion-induced adopters can be less valuable because promotional adoption is associated with forward buying and lower subsequent spend. Thus, managers should not assume that customers induced to adopt online shopping through external events or promotions will behave like *organic adopters* after adoption.

Our results have important implications for two core managerial tasks: forecasting Customer Lifetime Value (CLV) and evaluating the ROI of promotions. We show that ignoring adoption pathways can lead to substantial bias in incremental CLV forecasts. In our setting, pooled-adopter CLV forecasts substantially overstate the value of promotion-induced adopters: pathway-aware CLV is about 60% lower than the pooled forecast for *Black Friday adopters* and about 80% lower for *Loyalty Program adopters*. Similarly, using *organic adopters* as the benchmark in promotion evaluation makes online-adoption promotions appear to break even three to six times faster than they actually do. These results imply that firms should explicitly account for adoption pathways when forecasting incremental CLV, designing promotions, and setting caps on acceptable acquisition or promotional costs.

Overall, our analysis makes both substantive and managerial contributions. Substantively, we extend the multichannel shopping literature by introducing channel adoption pathways as a key dimension of heterogeneity and by documenting how organic, macro-shock-driven, and promotion-driven adopters follow systematically different post-adoption trajectories in spend, channel mix, and profitability. Managerially, we provide a cautionary tale against naively extrapolating the behavior and profitability of one adopter group (e.g., *organic adopters*) to other groups, and offer a framework for incorporating adoption pathways into CLV forecasting, promotion design, and ROI/break-even analysis.

2.2 Related Literature

Table 2.2. Prior Empirical Literature on Multichannel Shopping in Retail Settings

Paper	Setting	Channels	Methodology	Adoption Pathways
Our Paper	Pet Supplies	Physical stores only to adding website and/or mobile app	Diff-in-diff	Organic + multiple external stimuli
Narang and Shankar (2019)	Video games and Electronics	Physical stores + website to adding mobile app	Diff-in-diff	Organic only
Montaguti et al. (2016)	Books (subscribers only)	Mail order, fax, phone, SMS, Internet, and bookstores	Cross-sectional	Organic + mailing campaign
Wang et al. (2015)	Online grocery	Website to adding mobile app	Diff-in-diff	Organic only
Kushwaha and Shankar (2013)	Multiple categories	Website and catalog	Cross-sectional	Organic only
Ansari et al. (2008)	Apparel	Website and catalog	Within-customer	Organic only
Venkatesan and Ravishanker (2007)	Apparel	Full price stores, Discount Stores, and website	Within-customer	Organic only
Kumar and Venkatesan (2005)	Technology	Salespersons, direct mail, telephone and website	Cross-sectional	Organic only
Thomas and Sullivan (2005)	Retail	Catalog, physical stores, website	Cross-sectional	Organic only

Our work contributes to the empirical literature on multichannel customer management, which studies how customers' behavior changes when they purchase through multiple channels rather than a single channel. Table 2.2 summarizes closely related studies; for broader reviews see Neslin et al. (2006); Neslin and Shankar (2009); Verhoef (2012).

Substantively, the existing literature overlooks heterogeneity in channel adoption pathways, implicitly assuming that the value of multichannel customers is independent of the motivations underlying their adoption decisions. As Table 2.2 shows, almost all prior work focuses on *organic adopters*. Montaguti et al. (2016) is an exception, using a mailing campaign to nudge consumers into adopting multiple channels. However, they do not compare organic multichannel customers with those nudged by the campaign, whereas this comparison is central in our paper. Moreover, our setting is a fast-moving consumer packaged goods retailer (pet supplies), whereas much of the earlier literature examines books, apparel, video games, electronics, and technology. These differ from our context along several dimensions, including purchase frequency, competitive intensity, and product type (often more hedonic than utilitarian), which has been shown to matter for multichannel effects (Kushwaha and Shankar, 2013).

Thus, we extend the multichannel literature by (i) examining how post-adoption behavior varies with channel adoption pathways and providing theoretical explanations for these differences, and (ii) focusing on a practically important channel transition—physical store only to combined physical store and online usage—in a fast moving consumer packaged goods setting, that has not been separately studied.

From a methodological perspective, prior research has used three broad approaches to documenting the value of adopting new channels: cross-sectional, within-customer, and difference-in-differences (DID). Early work (Thomas and Sullivan, 2005; Kumar and Venkatesan, 2005; Kushwaha and Shankar, 2013; Montaguti et al., 2016) uses cross-sectional comparisons of multichannel versus single-channel customers and concludes that multichannel customers tend to be more valuable. However, cross-sectional comparisons lack individual-specific controls and cannot isolate how behavior *changes* after multichannel adoption. A second stream (Venkatesan and Ravishanker, 2007; Ansari et al., 2008) employs within-customer analysis, comparing a customer's purchase behavior before versus after adopting multiple channels. This work also finds higher spending after multichannel adoption, but without an appropriate control group of non-adopters, it requires strong assumptions that no other time-varying factors affect behavior. More recent work (Wang et al., 2015; Narang and Shankar, 2019) uses DID with non-adopters as a control group, addressing many of these concerns. Our paper also employs a DID framework with non-adopters as a control group, but brings this methodology to a new setting and focuses on heterogeneity across adoption pathways, which has not been explored in prior empirical work.

2.3 Background, Goals, and Theory

As discussed in §2.2, the existing multichannel literature has not considered heterogeneity in channel adoption pathways and has implicitly assumed that all multichannel customers behave similarly, regardless of why they adopted new channels to shop with a given retailer. Similarly, industry reports (PRNewswire, 2023; Deloitte, 2012) also discuss how multi-channel consumers spend more than single-channel consumers without considering the reasons for multi-channel adoption. However, consumers may adopt a retailer's online shopping channels for many different reasons, and it is plausible that these different pathways are associated with differences in post-adoption behavior. At a high level, there are three broad classes of online adoption pathways: (1) *organic adoption*, where customers voluntarily adopt a retailer's online shopping channels for the added convenience of an additional channel; (2) *macro-environmental shocks*, such as public health crises (e.g., COVID-19), nearby store closures, or severe weather events that prevent store visits; and (3) *firm-driven promotions*, where retailers explicitly incentivize adoption of their online channels.

Each pathway represents a different type of consumer self-selection, i.e., different types of consumers choose to adopt a retailer's online shopping channels in response to each stimulus. Our goal is not to rule out selection effects; rather, we seek to document whether and how different reasons for adopting online shopping with a retailer affect post-adoption behavior. From a managerial perspective, firms cannot randomize online adoption across consumers; they can only "encourage" adoption through promotional activities and observe how different encouragements translate into different adoption pathways and subsequent purchases.[1] Accordingly, we aim to document the ATT of online adoption on post-adoption behavior for

[1] We refer readers to the broader literature on "encouragement design" in marketing and the social sciences. Broadly, this literature distinguishes between interventions where firms can assign consumers to treatment arms (e.g., pricing, advertising, web page design) and those where consumers must self-select into the focal action (e.g., signing up for a newsletter, adopting online shopping) (Bradlow, 1998; Syrgkanis et al., 2019; Mummalaneni et al., 2023). In the latter, firms can only encourage the action (e.g., via promotions), and even a fully randomized experiment cannot eliminate selection because the outcome of interest (online adoption) is itself chosen by the consumer. In such settings, the natural estimand is the Average Treatment Effect on the Treated (ATT), rather than the Average Treatment Effect (ATE).

different pathways and to relate differences in these ATTs to theories of consumer behavior.

We focus on two behavioral mechanisms. The first is consumer inertia and habit persistence. Organic adoption likely signals a customer's readiness or preference for online shopping. In contrast, adoption induced by a shock or incentive may reflect a temporary disruption rather than a durable preference shift. Consumer inertia and habit theory (Verplanken and Wood, 2006; Wood and Neal, 2009) therefore suggest that externally induced adopters may use the newly adopted online channel less intensively after adoption and may revert more strongly toward prior offline routines. This mechanism yields two related predictions. First, externally induced adopters should allocate a larger share of post-adoption spending to the offline channel than *organic adopters*. Second, this gap should be larger when the external event attracts older customers, because older consumers tend to be more entrenched in existing habits and exhibit greater inertia (Yang and Ching, 2014). Differences in post-adoption channel mix may have profit implications for retailers. Profit margins can differ across online and offline channels because of differences in prices, costs, fulfillment, or promotional intensity. As a result, two adopter groups with similar total spend can differ in profitability if they allocate spending differently across channels. The direction of this effect depends on relative channel margins. In our empirical setting, offline margins are higher on average, so greater offline persistence can raise profitability even when total spend is similar.

The second mechanism is forward buying. Time-limited promotions can induce customers to stock up or shift purchases into the promotional period, reducing subsequent demand (Neslin et al., 1985; Hendel and Nevo, 2006; Van Heerde and Neslin, 2017). This mechanism is particularly relevant for promotions designed to induce online adoption. Customers may adopt online shopping to take advantage of a discount or offer, spend more during the adoption period, and then spend less afterward. Thus, promotion-induced adoption may generate lower post-adoption spending than organic adoption, even if the promotion successfully moves customers online.

Based on these mechanisms and previous literature on multichannel consumers, we seek to empirically test the following predictions:

- H1: Customers who adopt a firm's online channels and become multichannel shoppers will increase their spending post-adoption relative to *offline-only customers* (consistent with prior findings on multichannel shoppers; see §2.2).
- H2: Customers with different channel adoption pathways will exhibit different post-adoption behavior in terms of spend, profitability, and channel mix. These differences will arise from a combination of customer demographics, consumer habit theory, and forward buying. Specifically,
  - H2a: Consumers who adopt in response to time-limited promotions may increase spending during the promotional period but reduce spending thereafter, dampening (or reversing) the spending boost from multichannel adoption in the post-promotion period.
  - H2b: Externally driven adopters (vs. organic adopters) may be less likely to show the same level and growth in the share of spending on the newly adopted online channel, as predicted by consumer inertia and habit theory.
  - H2c: The difference in online share between *organic adopters* and externally driven adopters may be larger when external factors cause relatively older consumers to adopt.
  - H2d: Differences in online share may translate into lower (higher) profitability for *organic adopters* if online channels have lower (higher) margins relative to offline channels.

2.4 Setting and Data

Recall that the key goals of the paper are: (1) to empirically document the effect of adopting online shopping on post-adoption consumption behavior, (2) to test whether different channel adoption pathways lead to different post-adoption behaviors, and (3) to verify that the differences in post-adoption behavior can be predicted using theoretical underpinnings of purchase behavior in retail settings, so as to provide generalizable insights for managers and researchers. To accomplish these goals, we need data from a setting where we can observe not only when an *offline-only consumer* adopts online shopping but also the driver of this decision, i.e., the adoption pathway. Further, we need sufficient variants in channel adoption pathways to meaningfully connect the post-adoption behavior to theory. Therefore, we focus on a retail data source that satisfies these requirements.

Our data comes from one of the largest Brazilian pet supplies retailers, which sells a wide variety of products for pets, such as pet food and snacks, medicines/pharmacy items, hygiene and grooming products, and pet toys and accessories. Pet supplies is a large and competitive industry with about $250 billion in annual revenue globally (Fortune, 2024). The firm has both offline and online channels. The retailer operates more than 200 physical stores in Brazil, which serve as the offline channel, and has a website as well as a mobile app, which serve as the online channel. Moreover, our data span 2019-2024 during which there were three different events including a macro/environmental shock (COVID-19) and two different types of promotions designed to incentivize online shopping (discussed in more detail in §2.5) which led to many consumers adopting online shopping with the firm. Further, our data are quite rich, and we are able to observe when a consumer adopts online shopping (and thus can connect the timing of adoption to the channel adoption pathway) and both the pre- and post-adoption consumption behavior. For these reasons, it is an ideal setting for our study.

Our data includes transaction-level records of all individual customers from both online and offline channels during the 63-month period from January 1st, 2019, to March 31st, 2024. Each transaction record is uniquely identified by an order number with one or more purchased items. For each purchased item, we have information on the unit price, quantity purchased, brand, purchased channel (online/offline), subcategory, and category (e.g., food, hygiene).

The firm tracks customer purchases using a National Identification Number (similar to a Social Security Number), which is unique to each individual in Brazil. Customers who purchase online (either through the website or the app) must be logged in to their profile, and therefore, all online purchases can be traced back to the National ID. Customers are also asked for their National ID when making an in-store purchase; however, this is not mandatory. Discussions with the retailer revealed that while most offline transactions are also linked with a customer profile, about a quarter of in-store transactions are not linked to a National ID. This number has remained fairly steady during the entire observation period and was not affected by COVID-19 or other marketing events[2].

A potential concern with any channel analysis is that the results may reflect assortment differences across the two channels. To ensure that this is not the case, we exclude data from four product categories that were not available in both channels for the entirety of the observation period; see Table A.1 in Appendix §A.1 for a summary of these categories and their availability information. Together, sales from these categories account for only 0.63% of the total sales during the observation period; as such, we do not expect these to play any significant role in the findings.

2.5 Estimation Sample and Summary Statistics

We first discuss the sample construction process for the three focal events in §2.5.1, and present the details of the samples for each of the three focal events in §2.5.2, and then present some summary statistics of the three samples of interest in §2.5.3.

2.5.1 Sample Construction

We now discuss how we construct the samples used in our empirical analyses and the main outcome variables of interest.

Focal events that triggered different pathways of adoption: Since we seek to connect channel adoption pathway to post-adoption behavior, we consider three events that each led to a surge in online adoption: (1) a macro-environmental shock (COVID-19), (2) a short-term promotion (Black Friday sales), and (3) a long-term loyalty program. The first event was

[2]Because the number of offline transactions not linked to customer IDs remained similar through our observation period and exhibited no significant shifts in response to any of the events that spurred online adoption (including COVID-19), it is unlikely that this issue systematically biases our findings on the adoption of online shopping.

a macroeconomic shock outside the firm's control, whereas the latter two were firm-driven marketing activities.

Customer cohorts: For each focal event, we first define three customer cohorts – (a) *event adopters*, consumers who adopted online shopping during the event, (b) *organic adopters*, consumers who adopted organically (without external incentives) just prior to the event, and (c) *offline-only customers*, consumers, who did not adopt online shopping at all.

Time periods: Next, we establish the pre-adoption, adoption, and post-adoption periods for each focal event as follows:

- The pre-adoption period is defined as the common period during which all three groups of customers purchase only through physical stores (i.e., before the onset of the focal event).
- The adoption period is defined as the time during which adopters started purchasing online. Because each event has two sets of adopters, we have two adoption periods:
  - Adoption period for *event adopters*: This is the period during which the focal event took place, and induced customers to adopt online shopping. For instance, in the case of COVID-19, this is the first month and a half since the onset of COVID-19. Similarly, for the Black Friday event, the adoption period for *event adopters* is simply the day of the Black Friday sale.
  - Adoption period for *organic adopters*: Since *organic adopters* serve as a control group for *event adopters*, for each event, we choose consumers who organically adopted online shopping *just before* the onset of the focal event, and this period varies based on the event.
- The post-adoption period refers to the months *after* the adoption periods of both *event adopters* and *organic adopters*.

See Table 2.3 for an overview of these periods for each of the three events. In our estimation sample for each event, we only include data from the pre- and post-adoption periods for that event. We exclude the adoption periods for both *event* and *organic adopters* from the estimation sample because consumers are in flux during that time (where some adopters have adopted while others have not). Furthermore, by excluding these observations, we avoid capturing any abnormal purchase behaviors associated with the initial adoption of online shopping and the onset of specific events (i.e., COVID-19, Black Friday, Loyalty Program). As such, it allows us to cleanly identify the post-online-adoption difference in customers' purchase behaviors; see Gu and Kannan (2021) for a similar specification. We now summarize the choice of these three periods for each event.

- COVID-19 event: We use one year of data before the adoption period as the pre-adoption period (the calendar year of 2019). For the post-adoption period, we use all the data from May 2020 till April 2023.
- Black Friday event: We use eight months of data, starting from January 2019 to October 2020, as the pre-adoption period. For the post-adoption period, we use data from December 2019 (following the Black Friday event) until the next Black Friday.
- Loyalty Program event: We use the data from five months before the adoption period, and all the data available (from the end of the adoption period till March 2024) as the post-adoption period.

Table 2.3. Pre-adoption, Adoption, and Post-adoption Periods Under Each Event

	COVID-19	**Black Friday**	**Loyalty Program**
Pre-adoption	2019-01-01 – 2019-12-31	2019-01-01 – 2019-10-31	2023-01-01 – 2023-05-31
Adoption			
Organic	2020-01-01 – 2020-02-29	2019-11-14 – 2019-11-28	2023-06-01 – 2023-07-31
Event-Induced	2020-03-16 – 2020-04-30	2019-11-29	2023-08-01 – 2023-09-30
Post-adoption	2020-05-01 – 2023-04-30	2019-12-01 – 2020-10-31	2023-10-01 – 2024-03-31

We discuss each event and the definition of the three cohorts of consumers in detail below.

Outcome metrics: For all our analyses, we consider three metrics that quantify a consumer's shopping behavior in a given month: (1) Total Spend, (2) Profitability, and (3) Share of Offline Spend. Formally, we are interested in the Average Treatment Effect on

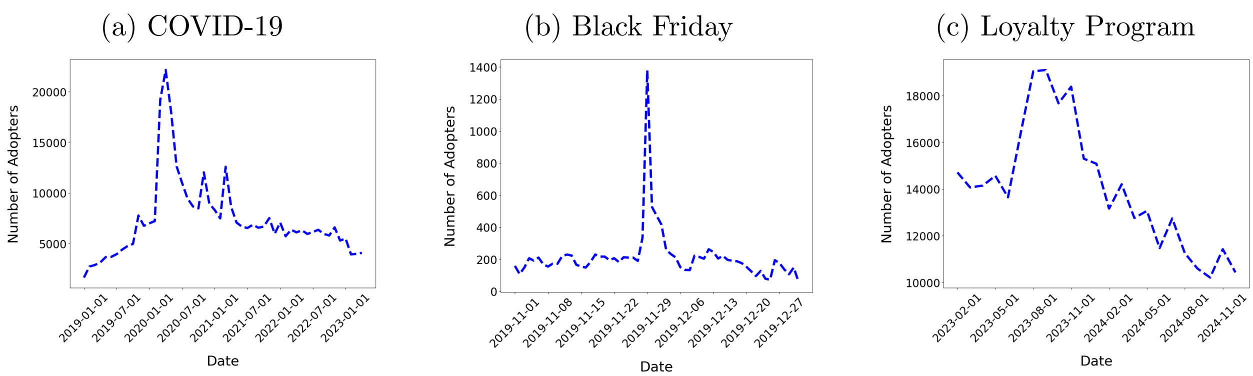
Treated (ATT) of adopting online shopping on these three metrics for each of the three focal events. Each of these metrics is defined at the customer-month level.

2.5.2 Estimation Samples for the Focal Events

We now describe the estimation sample used in the analysis of each of the three focal events.

2.5.2.1 Event 1: COVID-19

Figure 2.1. Number of *Offline-only* Customers First Adopting Online Shopping around the onset of COVID-19 (by month), Black Friday (by day), and Loyalty Program (by month)



The COVID-19 pandemic was an exogenous macro-environmental shock that prompted many consumers to adopt online shopping. Figure 2.1a shows the number of customers who adopted online shopping through the retailer's website and/or mobile app (from previously only shopping at physical stores) for each month for our observation period. As we can see, during March and April 2020, there was a large peak in the number of offline customers who first switched to online shopping; the number of adopters sharply increased from approx. 7,000 in January and February 2020 to approximately 20,000 in March and April 2020. Customers who adopted online shopping during the onset of COVID-19 may have done so for reasons like health concerns, social distancing norms, etc., whereas the customers who adopted online shopping before that did so of their own volition, for other reasons (e.g., the convenience of online shopping), unrelated to COVID-19.

We consider the set of customers who made their first offline purchase before 2019-12-31 and were offline-only through the end of 2019 and continued shopping with the firm after the onset of COVID-19. Then, we categorize each consumer from this set into one of the following three groups (see second column in Table 2.3 for exact dates):

- *Covid adopters*: Customers who made their first online purchase during the onset of COVID-19, i.e., from 2020-03-16 to 2020-04-30. 2020-03-16 was the date on which the Brazilian government began implementing physical distancing and confinement measures in response to COVID-19 (Faria de Moura Villela et al., 2021). Thus, we use this date as the start of the onset of COVID-19 and define the customers who adopted online shopping in the first one-and-a-half months as *Covid adopters*.
- *Organic adopters*: Customers who adopted online shopping organically in the period just before the onset of COVID-19. They made their first online purchase between 2020-01-01 and 2020-02-29, i.e., the two months before the onset of COVID-19.
- *Offline-only customers*: Customers who are active through both the pre- and post-adoption periods (i.e., at least one purchase before 2019-12-31 and after 2020-05-01), and never make online purchases during the estimation period.

In total, we have 33,246 *covid adopters*, 12,799 *organic adopters*, and 573,934 *offline-only customers* for this covid analysis. Note that a large fraction of consumers were able to remain offline-only during COVID-19 because the focal retailer's stores were allowed to remain open, as pet supplies retailing was deemed an essential service by the Brazilian government.

2.5.2.2 Event 2: Black Friday Sales

Next, we discuss the retailer's Black Friday sales event in 2019, characterized by a one-day, time-limited flash sales event offering discounts exclusively for online purchases. Because of the exclusive online nature of the event, consumers interested in these deals were motivated to adopt online shopping. Figure 2.1b indicates a peak in the number of adopters on 2019-11-29, which is the date that Black Friday sales took place.

For this estimation sample, we consider the cohort of customers who made their first offline purchase before 2019-10-31 and continued to be the firm's customers after the Black Friday sales date. Then, we categorize each consumer from this cohort into one of the following three groups (see third column in Table 2.3 for exact dates):

- *Black Friday adopters*: Customers who made their first online purchase during the 2019 Black Friday sales.
- *Organic adopters*: Customers who adopted online shopping organically in the two weeks preceding the 2019 Black Friday sales.
- *Offline-only customers*: Customers who are active through both the pre- and post-adoption periods, and never make online purchases during the estimation period.

In total, we have 1380 *Black Friday adopters*, 3061 *organic adopters*, and 534,871 *offline-only customers* for the Black Friday analysis.

2.5.2.3 Event 3: Loyalty Program

In August and September 2023,[3] the retailer launched a loyalty program and extensively advertised it to customers through email and store marketing. Customers who wanted to join the LP had to use the mobile app (no offline options were available for activating the loyalty program). After completing the online activation, customers could access and activate monthly special offers and receive discounts on both *online* and *offline* purchases. Note that customers had to activate each month to participate in that month's special offers. Figure 2.1c presents the number of adopters by month since January 2023 and shows clear peaks in the number of adopters during August and September 2023, coinciding with the launch of the loyalty program. This suggests that online-only activation of the newly available loyalty program prompted a significant number of customers to adopt online shopping during the

[3]Pilot studies for the loyalty program ran in two states (Rio Grande do Sul and Paraná) as early as May. Therefore, we exclude these two states from the estimation sample for this event.

roll-out period.

For this event, we consider the cohort of customers who made their first offline purchase before 2022-12-31 and continued to be the firm's customers during the post-online-adoption period (i.e., after 2023-10-01). Then, we categorize each consumer from this cohort into one of the following groups (see last column of Table 2.3 for details):

- *Loyalty Program adopters*: Customers who made their first online purchase in August or September 2023, with their first online purchase occurring *after* they activated the loyalty program online. In such cases, we can attribute the online loyalty activation as a reason for online adoption, as the activation precedes the first online purchase and both events occur within a similar time frame.
- *Organic adopters*: Customers who made their first online purchase in June or July 2023, prior to the launch of the loyalty program. These consumers adopt online shopping organically, but may still activate loyalty offers after their online adoption. We will take this into account in our empirical analysis.
- *Offline-only customers*: Customers who are active through both the pre- and post-adoption periods (i.e., at least one purchase before 2022-12-31 and after 2023-10-01), and never make online purchases during the estimation period.

In total, we have 1567 *loyalty program adopters*, 16,746 *organic adopters*, and 396,396 *offline-only customers* for this analysis.

2.5.3 Summary Statistics

For each consumer in the estimation sample of any focal event, we aggregate their transaction-level data into customer-month-level panel data over the customer's entire purchase history within the relevant period (January 2019 to March 2024), starting from their initial purchase month and ending with the last observed purchase month. For example, if a customer purchased for the first time in our data in August 2019 and the last observed purchase is in July 2020, then the panel for this customer spans 12 months, from August 2019 to July

2020. If there are intermediate months in which a customer did not make any purchase, the values for the monthly purchase-related variables are set to zero.

Thus, for each event (Covid-19, Black Friday, and Loyalty Program), we generate customer-month-level panel data for the pre- and post-adoption periods. These three panel datasets form the basis of our empirical analysis going forward. Table 2.4 presents the mean statistics of demographic (age and household income) and pre-adoption purchase behaviors (spend, quantity, orders, brands, categories, etc.) for the three groups of *event adopters*, their corresponding *organic adopters*, and *offline-only* customers. We further present the detailed summary statistics on demographics (gender and location), pre- and post- and adoption purchase behaviors for the three groups of customers (*offline-only*, *organic adopters*, and *event adopters*) from each event (Covid-19, Black Friday, and Loyalty Program) in Appendix §A.2. We discuss some key takeaways here.

Across our three samples for COVID-19, Black Friday, and the Loyalty Program, we find consistent differences between the two groups of adopters (*organic adopters* and *event adopters*) and *offline-only customers*. For instance, the adopter groups are consistently younger, higher income, and have higher spend with the retailer pre-adoption, relative to *offline-only customers*.

However, there are some important differences between the two groups of adopters (*event adopters* and *organic adopters*) across all three events. For instance, in the case of COVID-19, *event/covid adopters* tend to be older, have higher incomes, and have been shopping with the retailer for a longer time compared to *organic adopters*. However, these two groups have similar pre-adoption monthly spending. In the case of Black Friday, *event/Black Fridayadopters* are younger, have been shopping with the retailer for a longer time, and have lower monthly spend with the retailer in the pre-adoption period relative to *organic adopters*. In contrast, in the Loyalty program case, *event adopters* tend to be slightly older, have been shopping with the retailer for a shorter period, and have a higher monthly spend with the retailer in the pre-adoption period compared to *organic adopters*. However, the differences in characteristics between the two groups of adopters for each event are much smaller in

Table 2.4. Customer-Level Pre-adoption Summary Statistics Across Adoption Events

Variable	Adopters (Mean)	Organic (Mean)	Diff	t-stats	p-value	Offline-only (Mean)
Panel A1: COVID Adoption — Demographics and Pre-adoption Characteristics						
Age	43.36	41.44	1.92	15.65	0.00	45.62
Monthly Household Income	2432.11	2200.69	231.41	12.30	0.00	2033.01
Tenure (Days)	245.12	215.63	29.49	23.99	0.00	239.70
Panel B1: COVID Adoption — Pre-adoption Purchase Behaviors						
AvgSpendPerMonth	463.46	453.26	10.20	1.80	0.07	284.38
AvgQuantitiesPerMonth	4.98	4.43	0.55	6.15	0.00	2.96
AvgOrdersPerMonth	0.95	0.95	-0.00	-0.38	0.71	0.74
AvgUniqueItemsPerMonth	2.92	2.89	0.03	0.97	0.33	1.91
AvgUniqueBrandsPerMonth	2.32	2.34	-0.02	-0.92	0.36	1.55
AvgUniqueSubcategoriesPerMonth	2.02	2.04	-0.02	-1.10	0.27	1.39
AvgUniqueCategoriesPerMonth	1.65	1.66	-0.01	-0.62	0.54	1.17
AvgProfitPerMonth	162.12	162.36	-0.24	-0.12	0.91	97.52
Panel A2: Black Friday Adoption — Demographics and Pre-adoption Characteristics						
Age	41.86	42.72	-0.86	-2.49	0.01	47.31
Monthly Household Income	2278.60	2254.31	24.29	0.42	0.68	2046.46
Tenure (Days)	203.25	188.93	14.32	4.51	0.00	204.45
Panel B2: Black Friday Adoption — Pre-adoption Purchase Behaviors						
AvgSpendPerMonth	484.56	472.50	12.06	0.67	0.50	291.87
AvgQuantitiesPerMonth	5.00	4.49	0.50	2.07	0.04	3.00
AvgOrdersPerMonth	1.00	0.99	0.02	0.55	0.58	0.76
AvgUniqueItemsPerMonth	3.06	2.95	0.10	0.99	0.32	1.93
AvgUniqueBrandsPerMonth	2.44	2.39	0.05	0.72	0.47	1.57
AvgUniqueSubcategoriesPerMonth	2.13	2.08	0.04	0.74	0.46	1.41
AvgUniqueCategoriesPerMonth	1.72	1.69	0.03	0.77	0.44	1.19
AvgProfitPerMonth	165.90	168.36	-2.46	-0.36	0.72	99.52
Panel A3: Loyalty Program Adoption — Demographics and Pre-adoption Characteristics						
Age	44.57	43.43	1.14	2.78	0.01	47.01
Monthly Household Income	1959.34	2072.83	-113.49	-2.24	0.02	1975.35
Tenure (Days)	786.58	827.54	-40.96	-3.09	0.00	915.96
Panel B3: Loyalty Program Adoption — Pre-adoption Purchase Behaviors						
AvgSpendPerMonth	576.42	444.70	131.71	8.49	0.00	332.57
AvgQuantitiesPerMonth	4.31	3.25	1.06	6.44	0.00	2.62
AvgOrdersPerMonth	1.04	0.83	0.20	9.64	0.00	0.70
AvgUniqueItemsPerMonth	2.55	2.05	0.50	7.96	0.00	1.64
AvgUniqueBrandsPerMonth	2.00	1.64	0.36	8.48	0.00	1.31
AvgUniqueSubcategoriesPerMonth	1.70	1.41	0.29	8.76	0.00	1.14
AvgUniqueCategoriesPerMonth	1.42	1.20	0.22	8.57	0.00	0.98
AvgProfitPerMonth	198.59	164.46	34.12	6.04	0.00	125.81

magnitude relative to the differences between adopters and non-adopters (i.e. *offline-only customers*).

2.6 Empirical Analysis and Results

Our goal is to estimate the ATT (Average Treatment Effect on the Treated) of online adoption on post-adoption purchase behavior for different types of events that trigger the adoption of online shopping, and then connect these ATTs estimates with the different adoption pathways. To estimate our ATTs of interest, we use the difference-in-differences approach. This approach allows us to account for unobserved heterogeneity across customers and time-specific shocks. For each event, we use customer-month-level panel data for the estimation sample corresponding to that event and include customer-month observations from the pre- and post-adoption periods. We exclude the adoption periods from the estimation for the reasons discussed in §2.5.1.

The rest of this section is organized as follows. First, in §2.6.1 we compare *adopters* and *offline-only customers*. Next, in §2.6.2, we compare *event adopters* and *organic adopters*. Finally, in §2.6.3, we discuss parallel pre-trend tests for our difference-in-differences specifications, and in cases where parallel pre-trends fail, we present HonestDID results to assess the sensitivity of our measured effect sizes to violations of pre-trends.

2.6.1 Adopters vs. offline-only customers

We first estimate the ATT of online adoption by comparing *adopters* with *offline-only customers* for each of the three events: COVID-19, Black Friday, and the Loyalty Program. In this section, we focus on two outcomes: total spend and profitability. We do not analyze the share of offline spend here because channel mix is mechanically determined for *offline-only customers*, who never adopt the online channel during the estimation period. We return to channel mix in §2.6.2, where both comparison groups consist of online adopters.

We estimate the following two-way fixed effects (TWFE) specification with customer fixed effects (γ_i) and time (month) fixed effects (η_t):

$$y_{it} = \gamma_i + \eta_t + \beta_0 \mathbb{I}\{Post_t\} \times \mathbb{I}\{Adopter_i\} + \boldsymbol{\beta}_1' \boldsymbol{x}_{it} + \epsilon_{it}. \tag{2.1}$$

Here, y_{it} is the monthly outcome of interest for customer i in month t. $\mathbb{I}\{Adopter_i\}$ equals 1 for customers who adopt online shopping, including both *organic adopters* and *event adopters*, and 0 for *offline-only customers*. $\mathbb{I}\{Post_t\}$ equals 1 in the post-adoption period and 0 in the pre-adoption period. The coefficient of interest, β_0, captures the average incremental post-adoption change in the outcome for *adopters*, relative to *offline-only customers*. For the Loyalty Program analysis, we also include time-varying controls $\boldsymbol{x}_{it}$ that capture customers' participation in the loyalty program, such as whether they activate an offer. We describe these controls in Appendix §A.3.1.

This specification addresses our first research question and tests H1: whether customers who adopt the firm's online channel increase their spending and profitability after adoption, relative to *offline-only customers*.

Spend. Table 2.5 presents the results with monthly spend as the dependent variable. Across all three events, the coefficient on *Adopter* × *Post* is positive and statistically significant. To interpret the magnitudes of the estimated ATTs, we scale each ATT by the pre-adoption monthly spend of adopters, using the corresponding estimates from the standard two-by-two DID specification reported in Appendix §A.3.2. The implied effect sizes are substantial: monthly spend increases by 30.6% for COVID-19 adopters, 16.7% for Black Friday adopters, and 18.6% for Loyalty Program adopters. Thus, online adoption is associated with economically meaningful increases in spend across all three events.

Profitability. We next examine whether the increase in spend translates into higher profitability. In our data, we observe item-level profit margins for each purchase, which allows us to construct monthly customer-level profitability. Table 2.6 presents the results with monthly profit as the dependent variable. The coefficient on *Adopter* × *Post* is again positive and statistically significant across all three events. To interpret the effect sizes, we scale

Table 2.5. DID Main Analysis – Spend (*Adopters* vs. *Offline-only*)

Study:	Covid	Black Friday	Loyalty Program
DV:		Spend	
Variables			
Adopter × Post	140.1 (2.974)	78.37 (7.078)	84.89 (4.238)
	[0.000]	[0.000]	[0.000]
% Effect size relative to the pre-adoption baseline	30.6%	16.7%	18.6%
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	22,204,608	9,505,999	4,403,449
R^2	0.38214	0.41557	0.48127
Within R^2	0.00102	6.84×10^{-5}	0.00294

Clustered (Customer) standard errors in parentheses; p-values in brackets.

Notes: Percentage effects equal the DID estimate divided by the pre-adoption mean of adopters, reported in Table A.11 in Appendix A.3.2.

the estimates by adopters' pre-adoption monthly profitability. We get effect sizes of 24.3%, 11.7%, and 11.0%, respectively. Thus, the post-adoption increase in spend is accompanied by a corresponding increase in profitability.

Table 2.6. DID Main Analysis – Profit (*Adopters* vs. *Offline-only*)

Study:	Covid	Black Friday	Loyalty Program
DV:		Profit	
Variables			
Adopter × Post	37.37 (1.070)	18.44 (2.506)	18.41 (1.506)
	[0.000]	[0.000]	[0.000]
% Effect size relative to the pre-adoption baseline	24.3%	11.7%	11.0%
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	22,204,608	9,505,999	4,403,449
R^2	0.31903	0.34750	0.44546
Within R^2	0.00047	2.48×10^{-5}	0.00193

Clustered (Customer) standard errors in parentheses; p-values in brackets.

Notes: Percentage effects equal the DID estimate divided by the pre-adoption mean of adopters, reported in Table A.12 in Appendix A.3.2.

Summary of key takeaways. The results in this subsection support H1. Across all three events, customers who adopt online shopping subsequently spend more and generate higher profit than comparable *offline-only customers*. The magnitude of these effects is economically meaningful: spend increases by 16.7% to 30.6%, while profitability increases by 11.0% to 24.3%, depending on the adoption setting.

These results are consistent with prior work showing that multichannel customers are more valuable than single-channel customers (Neslin and Shankar, 2009; Montaguti et al., 2016). However, this comparison pools *organic adopters* and *event adopters* into a single adopter group. As a result, it abstracts away from heterogeneity in why customers adopted online shopping. In the next subsection, we examine this heterogeneity directly by comparing each group of *event adopters* with *organic adopters*.

2.6.2 Event adopters vs. organic adopters

We next examine whether post-adoption behavior differs by the reason for online adoption. Specifically, for each event, we compare the corresponding group of *event adopters* with *organic adopters*. This comparison isolates heterogeneity across adoption pathways among customers who all adopt online shopping. We estimate the following two-way fixed effects (TWFE) specification:

$$y_{it} = \gamma_i + \eta_t + \beta_2 \mathbb{I}\{Event\_Adopter_i\} \times \mathbb{I}\{Post_t\} + \boldsymbol{\beta}_3' \boldsymbol{x}_{it} + \epsilon_{it}. \tag{2.2}$$

Here, y_{it} is the monthly outcome of interest for customer i in month t. γ_i and η_t denote customer and month fixed effects, respectively. $\mathbb{I}\{Event\_Adopter_i\}$ equals 1 for the relevant group of *event adopters* and 0 for the corresponding group of *organic adopters*. $\mathbb{I}\{Post_t\}$ equals 1 in the post-adoption period and 0 in the pre-adoption period. The coefficient of interest, β_2, captures the difference in the post-adoption change in the outcome between *event adopters* and *organic adopters*. For the Loyalty Program analysis, we include the same time-varying loyalty-program controls described in Appendix §A.3.1.

We report the TWFE estimates as our main estimates. We also estimate a standard two-by-two DID specification for all the outcomes of interest and report those results in Appendix §A.4.1. The two-by-two DID results are consistent with the TWFE results and are used to construct the effect-size benchmarks reported in the main tables.

To facilitate interpretation, we report effect sizes using a common set of conventions throughout this subsection. For spend and profitability, we report two benchmarks. First, we scale the DID estimate by the post-adoption change for *organic adopters*, measured by the *Post* coefficient in the corresponding two-by-two DID specification. This benchmark asks whether the difference between *event adopters* and *organic adopters* is large relative to the organic adopters' own post-adoption change. If this denominator is not statistically significant, we report the relative effect as NA because the comparison is not well defined. Second, we scale the DID estimate by the pre-adoption mean of *event adopters*, which gives the magnitude of the differential effect relative to the event-adopter group's own baseline. If the DID estimate itself is statistically insignificant, we report the effect size as 0.0% and interpret the difference as statistically indistinguishable from zero. When both the DID estimate and the organic-adopter post-adoption change are negative, a positive value under the first benchmark indicates an additional decline relative to the organic-adopter decline, not a positive effect. For share of offline spend, the DID coefficient is already interpretable as an incremental difference in share relative to *organic adopters*; we report this effect in percentage points. We do not scale share of offline spend by the pre-adoption baseline because all customers are offline-only before adoption, so the pre-adoption offline share is mechanically 100% for all positive-spend customer-months.

We first present results for spend in §2.6.2.1, then examine the share of offline spend in §2.6.2.2, and finally analyze profitability in §2.6.2.3. These analyses address our second research question and test the second set of predictions, H2a through H2d, presented in §2.3.

2.6.2.1 Spend

Main DID results and economic magnitudes Table 2.7 presents the results from estimating Equation (2.2) with monthly spend as the dependent variable. The coefficient on $Event\_Adopter \times Post$ is small and statistically insignificant for *covid adopters*, indicating no detectable difference in post-adoption spend between *covid adopters* and *organic adopters*. In contrast, the estimates for both promotion-driven adoption pathways are negative and statistically significant.

The effect-size rows in Table 2.7 show that the promotion-driven differences are economically meaningful. Relative to their own pre-adoption baselines, *Black Friday adopters* spend 14.2% less and *Loyalty Program adopters* spend 9.2% less than comparable *organic adopters* in the post-adoption period. The relative-to-organic benchmark also shows that the Loyalty

Table 2.7. DID Main Analysis – Spend (*Event Adopters* vs. *Organic Adopters*)

Study:	Covid	Black Friday	Loyalty Program
DV:		Spend	
Variables			
Event_Adopter × Post	-3.095 (6.465)	-69.88 (14.26)	-53.22 (19.42)
	[0.632]	[0.000]	[0.006]
% Effect size relative to the post-adoption change for organic adopters	0.0%	NA	177.9%
% Effect size relative to the pre-adoption baseline for event adopters	0.0%	-14.2%	-9.2%
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	1,783,949	76,842	197,658
R^2	0.39193	0.41147	0.47126
Within R^2	5.37×10^{-7}	0.00068	0.02692

Clustered (Customer) standard errors in parentheses; p-values in brackets.

Notes: Effect sizes are calculated using the standard two-by-two DID estimates reported in Table A.13 in Appendix A.4.1. The effect-size conventions are described at the beginning of §2.6.2.

Program reduction is large: the additional decline for *Loyalty Program adopters* is 177.9% of the post-adoption decline observed among *organic adopters*. For *covid adopters*, the spend difference is statistically indistinguishable from zero.

Taken together, these results show that the adoption pathway matters for post-adoption spend. Customers who adopt online shopping during COVID-19 spend similarly to *organic adopters* after adoption. By contrast, customers who adopt through firm-driven promotions spend less than *organic adopters* after adoption. This pattern is consistent with H2a and with theories of forward buying and demand distortion (Neslin et al., 1985; Hendel and Nevo, 2006; Van Heerde and Neslin, 2017): time-limited promotions can induce customers to pull future demand into the promotional period, reducing spending in the post-promotion period.

Forward buying mechanism. The negative post-adoption spend effects for *Black Friday adopters* and *Loyalty Program adopters* may arise because these customers increase spending during the adoption period itself and then reduce spending afterward. That is, promotion-driven adopters may pull forward future demand, generating a post-promotion dip in spending.

To examine this mechanism, we estimate an event study specification that compares each promotion-driven adopter group with the corresponding group of *organic adopters* in each month around adoption:

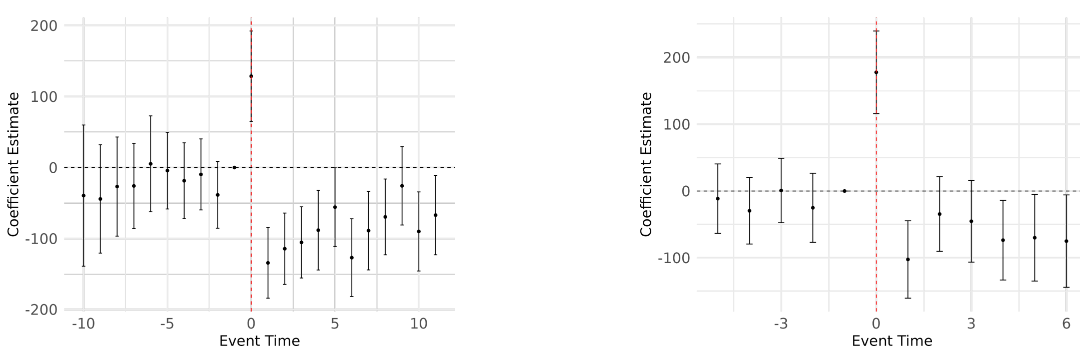
$$\begin{aligned} y_{it} = \gamma_i + \eta_t + \sum_j \theta_j \left(PRE_{it}(j) \cdot \mathbb{I}\{Event\_Adopter_i\} \right) \\ + \sum_k \theta_k \left(POST_{it}(k) \cdot \mathbb{I}\{Event\_Adopter_i\} \right) + \boldsymbol{\lambda}' \boldsymbol{x}_{it} + \epsilon_{it}. \end{aligned} \tag{2.3}$$

Here, $PRE_{it}(j)$ is an indicator that equals 1 if month t is j months before the adoption month. The coefficients on these pre-adoption indicators allow us to assess differential pre-trends between *event adopters* and *organic adopters*; we discuss these parallel-trend diagnostics in §2.6.3. Similarly, $POST_{it}(k)$ is an indicator that equals 1 if month t is k months after adoption. The corresponding coefficients capture the dynamic differences in spend between *event adopters* and *organic adopters* around and after adoption. We normalize outcomes relative to the month immediately prior to adoption.

We estimate Equation (2.3) for Black Friday and the Loyalty Program and report the full estimation results in Appendix §A.4.1. Figure 2.2 plots the dynamic treatment effects and 95% confidence intervals. The evidence is consistent with forward buying for both promotion-driven adoption pathways. During the month of adoption, *Black Friday adopters* and *Loyalty Program adopters* spend more than *organic adopters*. In the months following adoption, they spend less.[4] We also confirm that for COVID-19, which is not a firm-driven promotion, there is no comparable evidence of forward buying; see Figure A.2 in Appendix §A.4.1.

Figure 2.2. Event Study for Forward Buying: Both *Black Friday adopters* and *Loyalty Program adopters* spend more than *organic adopters* during the month of adoption. However, they spend less than *organic adopters* in the following months

(a) Black Friday (b) Loyalty Program



Summary of key takeaways. The spend results provide three main takeaways. First, *covid adopters* do not differ meaningfully from *organic adopters* in post-adoption spend. Second, promotion-driven adopters spend less than *organic adopters* after adoption, with

[4] Notably, when we examine offline spending only, there are no significant differences in offline spending between *Black Friday adopters* and *organic adopters*; see Figure A.1a in Appendix §A.4.1. This suggests that the forward buying effect for Black Friday is specific to the online channel. For the Loyalty Program, we observe the effect across both channels.

economically meaningful reductions of 14.2% for *Black Friday adopters* and 9.2% for *Loyalty Program adopters*, relative to their respective pre-adoption baselines. Third, the event study evidence shows that these post-adoption reductions are consistent with forward buying: promotion-driven adopters spend more during the adoption period but less afterward. Thus, managers should not expect promotion-driven online adopters to generate the same post-adoption spending gains as *organic adopters*.

2.6.2.2 Share of offline spend

We next examine how adoption pathways affect post-adoption channel usage. As discussed in §2.3, consumer inertia and habit theory (Verplanken and Wood, 2006; Wood and Neal, 2009) predict that adoption induced by external shocks or incentives may temporarily disrupt existing shopping routines, but that consumers may revert toward prior habits absent continued reinforcement. In our setting, this means that customers who adopt online shopping because of COVID-19, Black Friday, or the Loyalty Program may continue allocating a larger share of their spending to the offline channel than customers who adopt online shopping organically. This analysis tests H2b by comparing the post-adoption share of offline spend for each group of *event adopters* relative to the corresponding group of *organic adopters*.

The outcome variable is the share of monthly spending that occurs offline. It is defined only for customer-months with positive total spend; customer-months with zero total spend are treated as missing because the offline spending share is undefined in those months.

Main DID results and economic magnitude. Table 2.8 presents the results from estimating Equation (2.2) with the share of offline spend as the dependent variable. Since this outcome is a share, the coefficient on $Event\_Adopter \times Post$ is directly interpretable as the incremental post-adoption difference in offline spending share for *event adopters* relative to *organic adopters*. The estimates imply that *covid adopters* allocate 6.98 percentage points more of their post-adoption spending to the offline channel, while *Black Friday adopters* allocate 9.22 percentage points more. Both differences are statistically significant. In contrast, the estimate for *Loyalty Program adopters* is only 0.18 percentage points and is statistically insignificant, indicating no detectable difference in post-adoption channel mix relative to *organic adopters*.

Table 2.8. DID Main Analysis – Share of Offline Spend (*Event Adopters* vs. *Organic Adopters*)

Study:	Covid	Black Friday	Loyalty Program
DV:	Share of Offline Spend		
Variables			
Event_Adopter × Post	0.0698 (0.0042)	0.0922 (0.0120)	0.0018 (0.0114)
	[0.000]	[0.000]	[0.889]
Incremental difference in share of offline spend relative to *organic adopters* (percentage points)	6.98	9.22	0.18
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	949,448	42,299	110,339
R^2	0.50739	0.53685	0.57059
Within R^2	0.00140	0.00528	0.00904

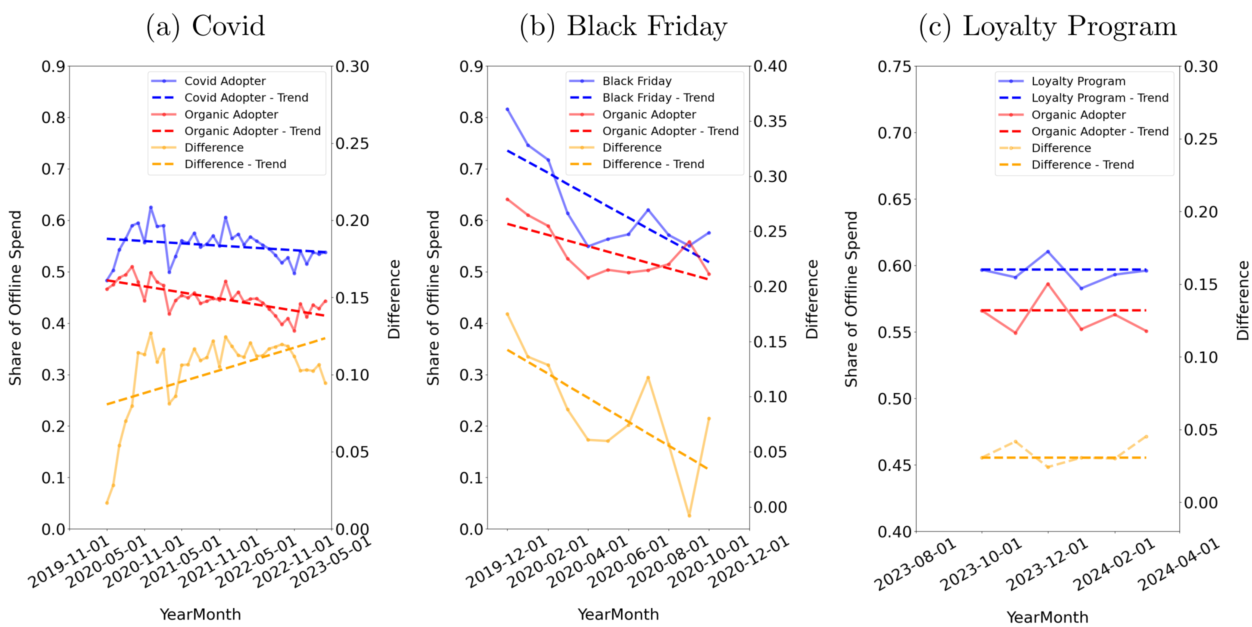
Clustered (Customer) standard errors in parentheses; p-values in brackets.

Notes: The percentage-point row equals the DID coefficient multiplied by 100. The effect-size conventions are described at the beginning of §2.6.2.

Dynamic patterns and mechanism. The DID estimates above summarize average differences over the full post-adoption period. We next examine whether these differences persist, widen, or narrow over time. Figure 2.1 plots the share of offline spend for each group of *event adopters* and the corresponding group of *organic adopters*, along with the difference between the two groups.

For COVID-19, Figure 2.3a shows that *covid adopters* not only have a higher post-adoption offline spending share than *organic adopters*, but also move toward online shopping more slowly over time; see Appendix §A.4.2 for formal tests of the magnitude of the difference. This pattern is consistent with consumer inertia and habit theory, and supports H2b.

Figure 2.3. Share of Offline Spend Relative to Total Spend – *Event Adopters* vs. *Organic Adopters*



Theory and mechanism. The COVID-19 pattern is also consistent with H2c. As shown in §2.5.3, *covid adopters* are older than *organic adopters*. Older consumers are more likely to be set in their shopping habits and exhibit greater inertia (Yang and Ching, 2014). To examine whether the slower movement toward online shopping is concentrated among older customers, we split *covid adopters* and *organic adopters* by median age. We find that the difference in offline-share trends appears among older adopters, but not among younger adopters.[5] This supports the interpretation that the channel-mix differences for *covid adopters* are re-

[5]See Figure A.3 and Table A.18 in Appendix §A.4.2 for details.

lated to consumer inertia and age-based differences in habit persistence.

For Black Friday, Figure 2.3b shows that *Black Friday adopters* initially allocate a higher share of spending to the offline channel than *organic adopters*, consistent with the positive DID estimate in Table 2.8. However, they also increase their online share of spend faster over time.[6] This pattern may reflect the fact that *Black Friday adopters* are younger than *organic adopters*, and thus may be less entrenched in offline habits and more willing to explore the convenience of online shopping. Nevertheless, because their initial adoption was induced by an external incentive, they still begin the post-adoption period with greater offline reliance than *organic adopters*.

For the Loyalty Program, the DID estimate in Table 2.8 is small and statistically insignificant, and Figure 2.3c shows little separation between *Loyalty Program adopters* and *organic adopters*. This may be because the two groups are relatively similar in demographics and because the Loyalty Program requires recurring monthly online activation, which may reinforce online channel usage after adoption.

Summary of key takeaways. The share-of-offline-spend results show that externally induced adoption does not always produce the same post-adoption channel mix as organic adoption. *Covid adopters* and *Black Friday adopters* allocate significantly more of their post-adoption spending to the offline channel than *organic adopters*, by 6.98 and 9.22 percentage points, respectively. In contrast, *Loyalty Program adopters* do not differ significantly from *organic adopters* in their offline spending share. The dynamic patterns further suggest that age and reinforcement matter: older *covid adopters* move online more slowly, while younger *Black Friday adopters* appear to catch up over time, and *Loyalty Program adopters* show no meaningful channel-mix divergence, possibly because repeated app-based activation reinforces online use. These findings support the role of consumer inertia and habit persistence in shaping post-adoption channel usage.

[6] The p-value for the trend difference is 0.109; see Table A.17 in Appendix §A.4.2. Extending the post-adoption period further reduces the p-value.

2.6.2.3 Profitability

We next examine how adoption pathways affect profitability. The spend results in §2.6.2.1 show that promotion-driven adopters spend less than *organic adopters* after adoption, while *covid adopters* spend similarly to *organic adopters*. The channel-mix results in §2.6.2.2 show that *covid adopters* and *Black Friday adopters* allocate a larger share of their post-adoption spending to the offline channel. Profitability can therefore differ across adoption pathways for two reasons: differences in total spend and differences in the allocation of spend across channels.

Channel margin differences. To interpret the profitability results, we first document differences in prices and profit margins across the offline and online channels. We construct a panel at the item-channel-month level.[7] We then calculate offline-online differences at the item-month level and average these differences over time to obtain item-level average differences in price and profit margin across channels. Table 2.9 reports item-level summary statistics for the 22,803 items that were sold in both offline and online channels in our data and were purchased at least once in the same month. For confidentiality, we multiply profit margins by the same undisclosed multiplier used for prices.

Table 2.9. Item-Level Offline Price and Profit Margin Relative to Online

	mean	std	25%	50%	75%	(min, max)	count
Offline - Online Price	6.01	91.07	-0.84	4.61	17.71	(-7,598.46, 1,430.91)	22,803
Offline - Online Profit Margin	5.37	78.49	-1.77	3.47	14.05	(-6,586.62, 2,163.31)	22,803

On average, prices are higher in physical stores than online, and profit margins are also higher in physical stores than online.[8] Thus, even when two adopter groups have similar

[7]For instance, in the case of price, an observation in this panel is the average price of item j sold through channel c in month t, calculated as the total sales of item j in channel c during month t divided by the total quantity of that item sold in that channel-month.

[8]While the retailer in our context has higher offline prices, there is considerable heterogeneity across retailers, countries, and product categories in whether online prices are lower, the same, or higher than offline prices. In the US, 69% of multichannel retailers have identical prices online and offline, while 22%

total spend, the group with a higher share of offline spend may generate higher profit. This is the mechanism underlying H2d. At the same time, the retailer still has an incentive to promote online adoption because, as shown in §2.6.1, multichannel adoption is associated with higher total spending and profitability relative to remaining offline-only.

Table 2.10. DID Main Analysis – Profit (*Event Adopters* vs. *Organic Adopters*)

Study:	Covid	Black Friday	Loyalty Program
DV:		Profit	
Variables			
Event_Adopter × Post	9.079 (2.274)	-16.04 (5.094)	-15.41 (6.835)
	[0.000]	[0.002]	[0.024]
% Effect size relative to the post-adoption change for organic adopters	44.8%	NA	73.2%
% Effect size relative to the pre-adoption baseline for event adopters	5.8%	-9.9%	-7.8%
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	1,783,949	76,842	197,658
R^2	0.29792	0.35806	0.44335
Within R^2	2.79×10^{-5}	0.00025	0.02248

Clustered (Customer) standard errors in parentheses; p-values in brackets.

Notes: Effect sizes are calculated using the standard two-by-two DID estimates reported in Table A.19 in Appendix A.4.3. The effect-size conventions are described at the beginning of §2.6.2.

Main DID results and economic magnitude. Table 2.10 presents the results from estimating Equation (2.2) with monthly profit as the dependent variable. The coefficient on

have higher offline prices and only 8% have higher online prices (Cavallo, 2017).

Event_Adopter $\times$ *Post* is positive and statistically significant for *covid adopters*, indicating that they become more profitable than *organic adopters* after adoption. In contrast, the coefficients are negative and statistically significant for both *Black Friday adopters* and *Loyalty Program adopters*, indicating lower post-adoption profitability relative to *organic adopters*.

The effect-size rows in Table 2.10 show that these profitability differences are economically meaningful. For *covid adopters*, monthly profitability increases by 5.8% relative to their own pre-adoption baseline, and the incremental profitability gain is 44.8% of the post-adoption gain observed among *organic adopters*. This profitability advantage occurs despite similar total spend and is consistent with *covid adopters*' higher offline spending share, combined with the higher offline margins documented above. For *Black Friday adopters* and *Loyalty Program adopters*, profitability is lower than for *organic adopters*. Relative to their own pre-adoption baselines, *Black Friday adopters* generate 9.9% lower monthly profit and *Loyalty Program adopters* generate 7.8% lower monthly profit. For *Loyalty Program adopters*, the relative-to-organic benchmark also shows that the additional decline is sizable, amounting to 73.2% of the post-adoption decline observed among *organic adopters*.

Summary of key takeaways. The profitability results show that adoption pathways affect customer value in ways that are not captured by spend alone. *Covid adopters* are more profitable than *organic adopters* despite similar spending, consistent with their higher offline spending share and the higher offline margins in our setting. In contrast, *Black Friday adopters* and *Loyalty Program adopters* are less profitable than *organic adopters*, largely because they spend less after adoption, as shown in §2.6.2.1. These findings support H2d and show that firms should not assume that customers induced to adopt online shopping through external events or promotions will have the same post-adoption profitability as *organic adopters*.[9]

[9] We use Wald tests to confirm that the effect sizes for spend and profit for both promotion pathways, Black Friday and the Loyalty Program, are statistically indistinguishable from each other, while both are statistically different from the COVID adoption pathway.

2.6.3 Parallel pre-trends & HonestDiD Sensitivity Analysis

We test whether the parallel trends assumption holds for our DID analyses comparing spend and profitability between *event adopters* and *organic adopters*, and between adopters and *offline-only customers*, using both visual inspection and more formal event study regressions (Autor, 2003). The parallel trends assumption appears to be satisfied for both spend and profit when comparing *event adopters* and *organic adopters* for all three events. However, the parallel pre-trends assumption is not satisfied for the comparisons between adopters and *offline-only customers*, across all three events. Further, specifically for the case of *covid adopters* and *organic adopters*, there seems to be a difference in pre-trends visually, although the differences in pre-adoption outcomes are not statistically different for most of the pre-adoption months. Taken together, these patterns point to four settings in which the parallel trends assumption is less well supported: the comparisons of *adopters* versus *offline-only* customers across three events, and the comparison of *covid adopters* versus *organic adopters*. Figures 2.4 presents COVID-19 as a representative case for the total spend (top row) and profit (bottom row), while the full set of diagnostics for other two events (Black Friday and Loyalty Program) are reported in Appendix §A.5.

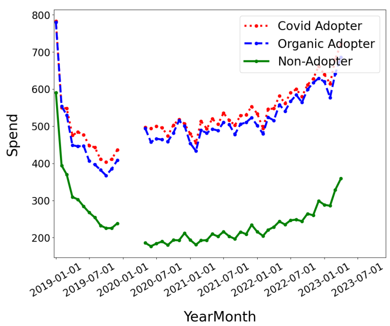
For these four cases (spend and profits for *adopters* vs. *offline-only customers* for all three events, and *covid adopters* vs. *organic adopters*), we bound our treatment effects under different assumptions on the extent of the violation using the HonestDiD approach proposed by Rambachan and Roth (2023). To conduct this HonestDiD sensitivity analysis, we impose the restriction that the change in slope of the differential trend between treatment and control customers across two periods is bounded by:

$$\Delta^{SD} := \{\zeta : |(\zeta_{t+1} - \zeta_t) - (\zeta_t - \zeta_{t-1})| \leq M,\ \forall t\} \tag{2.4}$$

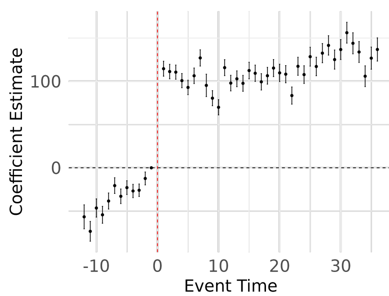
Here, ζ_t is the difference in trends between two groups at time t, and M is the magnitude of maximum allowable change in trend between consecutive periods. We test a wide range of M values, up to and beyond the linear pre-trend estimate (see Appendix §A.5.3 for the pre-trend estimates).

Figure 2.4. Parallel Trend Assessment – COVID-19. Panels (2.4a, 2.4d) plot monthly outcomes, while panels (2.4b, 2.4e) and (2.4c, 2.4f) report event study estimates (based on Equation (2.3)) comparing (i) *adopters* to *offline-only customers* and (ii) *covid* to *organic adopters*, respectively.

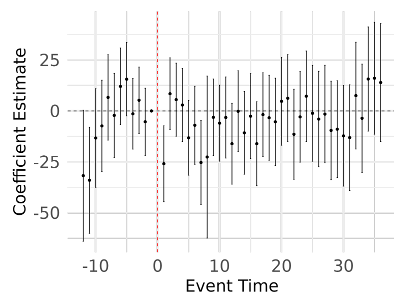
(a) Visualization – Spend



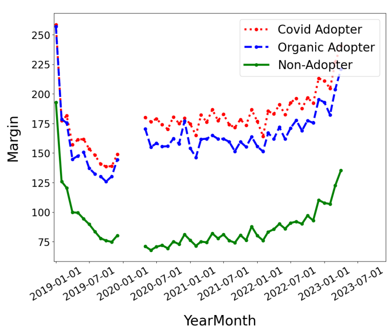
(b) *Adopters* vs. *Offline-only* – Spend



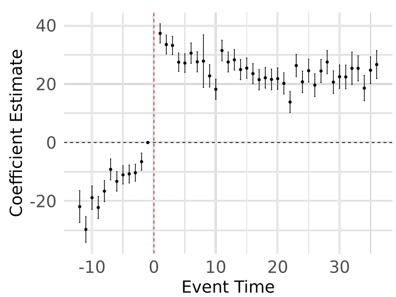
(c) *Covid* vs. *Organic Adopters* – Spend



(d) Visualization – Profit



(e) *Adopters* vs. *Offline-only* – Profit



(f) *Covid* vs. *Organic Adopters* – Profit

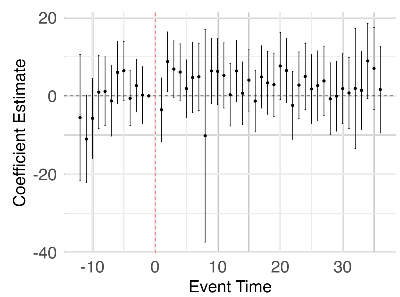
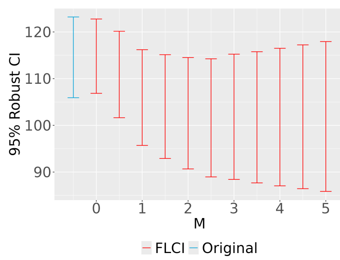


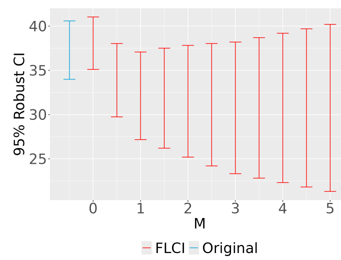
Figure 2.5. HonestDiD bounds across adoption pathways. Each panel reports HonestDiD 95% confidence interval bounds for the post-treatment effect (labeled as FLCI) under increasing restrictions (M) on deviations from parallel trends.

Panel A: Organic + Covid vs. Offline-only

(a) Spend

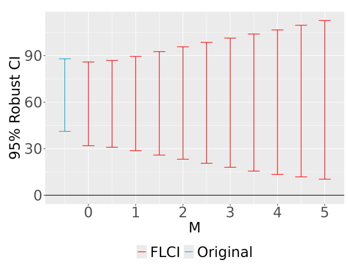


(b) Profit

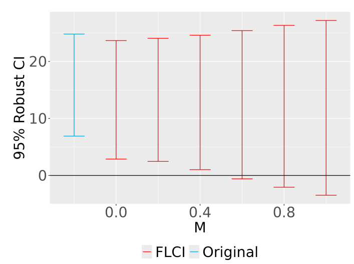


Panel B: Organic + Black Friday vs. Offline-only

(c) Spend

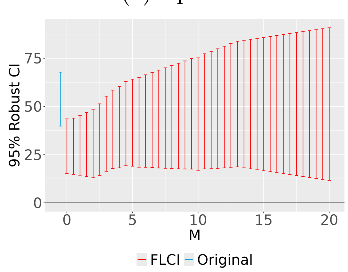


(d) Profit

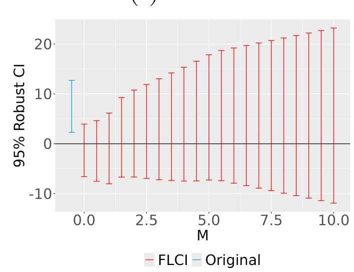


Panel C: Organic + Loyalty vs. Offline-only

(e) Spend

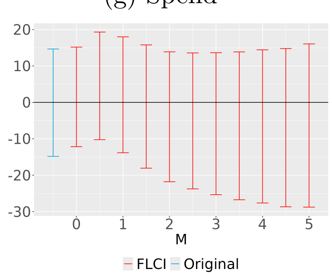


(f) Profit

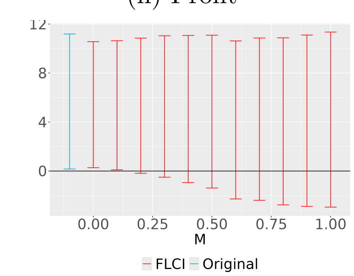


Panel D: Covid vs. Organic

(g) Spend



(h) Profit



We present the HonestDiD 95% confidence interval of ATT bounds under different violations M for spend and profit under four cases in Figure 2.5. Our results remain statistically significant for values for M larger than the estimated linear pre-trends. In all instances but one, we conclude that even if the post-adoption trend violation is twice as large as the pre-adoption trend, our substantive conclusions remain unchanged in terms of both direction and statistical significance. The only exception is profit for the Loyalty Program event. However, even for this event, the difference in post-adoption spend is large and significant.

2.7 Managerial Implications

Our main findings indicate that consumers who adopt online shopping through different channel pathways (organic, macro-environmental shocks, or promotions) behave differently in the post-adoption period. These patterns are well explained by consumer inertia and forward buying, and should generalize to other settings where similar mechanisms operate. In particular, forward buying and stockpiling can reduce the long-run increase in spending from promotion-induced online adopters. To mitigate this and maximize long-run value, firms may benefit from designing incentives that discourage stockpiling (e.g., phased or behavior-contingent rewards) and combine promotions with rapid follow-up nudges to encourage repeat (rather than one-off) behavior. For consumers pushed online by shocks (like COVID-19), habit-reinforcing promotions and personalized nudges in the immediate post-adoption window can help convert trials of online channels into stable habits, thereby improving long-run profitability. Beyond these general lessons, we derive concrete implications for two common marketing tasks: (a) CLV forecasting and (b) ROI/breakeven analysis for promotions.

2.7.1 Implications for CLV forecasts

Customer Lifetime Value (CLV) is a core metric for customer relationship management, targeting, and budgeting (Fader et al., 2005; Gupta et al., 2006; Fader et al., 2010). Our results show that when *offline-only customers* adopt online shopping for different reasons (e.g., Black Friday, COVID-19, loyalty program, or organically), their spending and profitability trajectories differ markedly by pathway. Ignoring this heterogeneity leads to biased forecasts of *incremental* CLV from online adoption and, in turn, to suboptimal decisions, for example, overestimating the effectiveness of promotions aimed at inducing online adoption.

We quantify the *average* incremental (relative to *offline-only customers*) CLV from online adoption for each adopter cohort s. Following standard practice (Gupta et al., 2006), the (average) incremental CLV equals the present value of all future incremental monthly profits attributable to adoption under an infinite horizon:

$$Incremental\ CLV_s = \sum_{t=1}^{\infty} \frac{ATT_s \times r_s^t}{(1+i)^t} = ATT_s \frac{r_s}{1+i-r_s}, \tag{2.5}$$

where: (1) ATT_s is the monthly average incremental profit (our ATT estimate) for cohort s relative to *offline-only customers*; (2) i is the firm's monthly cost of capital (WACC), and (3) r_s is the (average) monthly retention rate of cohort s post-adoption, defined as the probability of purchasing next month conditional on purchasing in the focal month. All quantities are monthly. The ATT_s estimates come from our DID models for each event, and retention rates are obtained non-parametrically. Further, based on discussions with the retailer, we set WACC to 15% annually, implying a monthly cost of capital of $i = (1+15\%)^{1/12} - 1 = 1.17\%$.

We next show how CLV forecasts can be biased when managers ignore adoption pathways. We consider two types of Incremental CLV estimates for each type of *event adopter*: (1) *Pooled* forecast, where the firm ignores the adoption pathway and pools *event* and *organic adopters* and estimates a single incremental CLV for all adopters, and (2) *Pathway aware* forecast estimates based only on the event adopters (without including *organic* adopters). The inputs as well as the results of this exercise are shown in Table 2.11. For the COVID-19 event, for example, the pooled ATT from Table 2.6 is 37.30 MCU, and the pooled monthly retention rate is 0.654. Overall, we see that pooled estimates overstate the value of promotion-induced adopters by almost 60% (Black Friday) and 80% (Loyalty Program), while understating the value of COVID-induced adopters by 7.5%.[10]

[10]Note that the difference for COVID-induced adopters is mainly a result of higher offline prices (vs.

Table 2.11. *Incremental* CLV (relative to *offline-only*, in MCU): Pathway-Aware (*Event Adopters*) vs. Pooled Forecasts – By Three Events (COVID-19, Black Friday, and Loyalty Program)

	COVID-19	Black Friday	Loyalty Program
Panel A: Inputs			
ATT (*Event Adopters*)	39.780	7.813	2.986
ATT (*Pooled*)	37.370	18.440	18.410
Monthly Retention Rate (*Event Adopters*)	0.656	0.621	0.625
Monthly Retention Rate (*Pooled*)	0.654	0.633	0.565
Panel B: Incremental CLV			
Incremental CLV (*Event Adopters*)	73.486	12.417	4.831
Incremental CLV (*Pooled*)	68.332	30.822	23.301
Panel C: Forecast Bias			
Diff in *Incremental* CLV of *Pathway aware* vs. *Pooled* Estimates	+7.5%	-59.7%	-79.3%

Notes: Incremental CLV is defined relative to the *offline-only* counterfactual and computed using Equation (2.5). ATT estimates for *event adopters* are reported in Table A.21 in Appendix §A.6. Pooled ATT estimates are from Table 2.6. Monthly retention rates for *event adopters* are estimated separately by adoption pathway, while pooled retention rates are estimated from a combined sample of *organic* and *event adopters*. Diff in *Incremental* CLV of *Pathway aware* vs. Pooled Estimates is defined as $\frac{(\text{Incremental}\quad \text{CLV\_pooled} - \text{Incremental}\quad \text{CLV\_event})}{\text{Incremental}\quad \text{CLV\_pooled}}$ for a given event.

Together, these findings suggest that failing to distinguish adoption pathways biases CLV forecasts—typically overstating the value of promotion-induced adopters and making weak promotional strategies appear attractive, which can lead to overspending and erode long-run profitability. A key takeaway is that retailers should (a) estimate (incremental) CLV separately by adoption pathway and (b) set pathway-specific caps on the cost of inducing online adoption, aligning offer depth and paid-media spend with those caps.

online) at our focal retailer. Not all retailers have higher offline prices. In the US, 69% of multi-channel retailers have identical prices online and offline, while 22% have higher offline prices and only 8% have higher online prices(Cavallo, 2017).

2.7.2 Implications for ROI and Breakeven Analyses for Promotions

We next examine how ignoring adoption pathways affects ROI and breakeven analyses for promotions designed to induce online adoption. We focus on the two firm-driven promotion pathways in our setting: Black Friday and the Loyalty Program. The key managerial question is whether the incremental profits generated by promotion-induced adopters are sufficient to recover the cost of the promotion. If managers use *organic adopters* as proxies for promotion-induced adopters, they may overstate the expected incremental profit from induced adoption and therefore understate the time required to break even.

For a promotion c and adopter benchmark group g, define $Cost_c$ as the per-customer cost of campaign c, and $ATT^{profit}_{c,g}$ as the average monthly incremental profit from online adoption for benchmark group g, relative to *offline-only customers*. The ROI over a T-month horizon is:

$$ROI_{c,g}(T) = \frac{T \times ATT^{profit}_{c,g} - Cost_c}{Cost_c}. \tag{2.6}$$

The breakeven time is the value of T that sets Equation (2.6) equal to zero:

$$Breakeven_{c,g} = \frac{Cost_c}{ATT^{profit}_{c,g}}. \tag{2.7}$$

We compare two benchmarks. The first is a naive benchmark that uses the incremental profit of *organic adopters*. This is the type of benchmark a manager might use when evaluating a new promotion using historical organic-adopter behavior. The second is a pathway-aware benchmark that uses the incremental profit of the actual promotion-induced adopters.

Table 2.12 reports the inputs and breakeven calculations. For Black Friday, we quantify the promotion cost as the difference between actual consumer spending at promotional prices and what consumers would have paid at regular prices. The average discount is about 28% across categories, yielding an average promotional cost of 299.18 MCU per customer. For the Loyalty Program, we use the retailer-provided cost assumption of 100 MCU per customer to set up the program and induce online activation.[11]

[11] This estimate is based on conversations with the retailer; we consider alternative values of 50 MCU and 150 MCU in Appendix §A.6.

Table 2.12. Breakeven Analysis for Promotions Designed to Induce Online Adoption

	Black Friday	Loyalty Program
Panel A: Inputs		
Campaign cost per induced adopter, $Cost_c$ (MCU)	299.18	100.00
Monthly incremental profit using *organic adopters* as benchmark, $ATT^{profit}_{c,organic}$ (MCU)	23.57	18.55
Monthly incremental profit using *event adopters* as benchmark, $ATT^{profit}_{c,event}$ (MCU)	7.813	2.986
Panel B: Breakeven Calculations		
Naive breakeven using *organic adopters*, $Cost_c/ATT^{profit}_{c,organic}$	299.18/23.57 = 12.69 months (≈ 13)	100/18.55 = 5.39 months
Pathway-aware breakeven using *event adopters*, $Cost_c/ATT^{profit}_{c,event}$	299.18/7.813 = 38.30 months (≈ 39)	100/2.986 = 33.49 months
Panel C: Managerial Implication		
Breakeven error from using *organic adopters* as proxies	Nearly 3× longer	More than 6× longer

Notes: Breakeven time is computed using Equation (2.7). Monthly incremental profit estimates for *organic adopters* and *event adopters* are from Table A.21 in Appendix §A.6. The Black Friday campaign cost is calculated as the difference between actual spending at promotional prices and counterfactual spending at regular prices. The Loyalty Program campaign cost is based on the retailer-provided cost estimate; alternative cost assumptions are reported in Appendix §A.6.

The results show that ignoring adoption pathways can substantially distort promotion evaluation. For Black Friday, using *organic adopters* as the benchmark implies a breakeven period of about 13 months. In contrast, using the incremental profit of actual *Black Friday adopters* implies a breakeven period of about 39 months, nearly three times longer. For the Loyalty Program, the distortion is even larger: the naive organic-adopter benchmark implies breakeven in 5.39 months, whereas the pathway-aware estimate implies breakeven in 33.49 months, more than six times longer.

In sum, managers must calibrate expectations of incremental profitability when running promotions to encourage online adoption. The key managerial error is not merely that promotion-induced adopters have different monthly profit effects; it is that using the wrong benchmark can make promotions appear to recover their costs far faster than they actually do. This can lead to overly optimistic ROI assessments, excessive promotional spending, and underestimation of the time required to recoup promotional investments.

2.8 Robustness Checks

We now present a series of robustness checks to provide additional credibility for our findings.

2.8.1 Effect size measurement

A potential concern with our ATT measurement, especially for our Covid analysis, is that *covid adopters* may be "contaminated" with some *organic adopters* – customers who would have adopted the online channel organically during the onset of COVID-19, had COVID-19 not occurred. As a result, the treatment group may include a mix of "true" *covid adopters* and *organic adopters*. If such contamination exists, the ATT estimates reported in §2.6.2 would be a lower bound on the true ATT for covid-driven adoption (relative to organic), and we can back out the ATT for the subset of customers of interest (i.e., the "true" *covid adopters*). The true ATT for *covid adopters* relative to *organic adopters* can be obtained by dividing the estimated ATT (reported in §2.6.2) by the fraction of "true" *covid adopters*, given by Equation (2.8):

$$\text{True } ATT_{covid,organic} = \frac{\text{Estimated } ATT_{covid,organic}}{w_{covid}} \tag{2.8}$$

where w_{covid} is the fraction of "true" *covid adopters* in the sample. Assuming the monthly incidence of *organic adopters* remains constant, we estimate that around 38.5% of the *covid adopters* could actually be *organic adopters*. Hence, $w_{covid} = 61.5\%$ can be the fraction of "true" *covid adopters*. Here, since both the Estimated $ATT_{covid,organic}$ and w_{covid} are estimated from the sample and thus contain sampling variability, we compute the standard errors of True $ATT_{covid,organic}$ using the block bootstrapping with 500 replications, which treat individual customers as the resampling units. In each replication, we resample the same number of customers (with replacement) as in the original data, re-estimate the DID, calculate the fraction of *covid adopters*, and obtain the bootstrapped standard errors for True $ATT_{covid,organic}$ (Cameron and Trivedi, 2005). Accordingly, the "true" ATT of spend for *covid-driven adopters* relative to *organic adopters* would be $\frac{-3.095}{0.615} = -5.033$ (std. error: 10.363, p-value: 0.627), and the "true" ATT of profit would be $\frac{-9.079}{0.615} = 14.763$ (std. error: 3.638, p-value: 0.000) [12]. This confirms that, even with potential contamination, the difference in spend relative to *organic adopters* is small, with a confidence interval that includes zero, whereas the difference in profit is larger and has a confidence interval that does not include zero. As such, our substantive conclusions remain unchanged.

Note that we are less concerned about the contamination problem for our other two events for a few reasons, which we briefly discuss below.

- First, since the Black Friday sale occurs only on a single day (and not over a prolonged period, unlike COVID-19), and there is a large spike in adopters on that single day compared to the days around it, the extent of potential contamination is very small. In this case, the results are very similar to those from §2.6.2 since historical data suggest that close to 85.7% of adopters during the Black Friday event adopt because of the event

[12] The ATT point estimates are taken from Tables 2.7 and 2.10. The standard errors are obtained from 500 bootstrap replications. The bootstrap distribution of True $ATT_{covid,organic}$ is presented in Figure A.6 in Appendix §A.7.1.

(and not organically). Nevertheless, we still conduct a similar analysis by backing out the "true" ATT using the estimated ATT of *Black Friday adopters* relative to *organic adopters* divided by $w_{black\_friday} = 85.7\%$. The "true" ATT of spend for *Black Friday adopters* relative to *organic adopters* would be $\frac{-69.88}{0.857} = -81.54$ (std. error: 20.236, p-value: 0.000), and the ATT of profit would be $\frac{-16.04}{0.857} = -18.716$ (std. error: 6.775, p-value: 0.006).[13] This confirms that even with such potential contamination, *Black Friday adopters* have lower spend and profit relative to *organic adopters*.

- Next, recall that the *Loyalty Program adopters* are customers who specifically went online for the first time to activate the loyalty program on their mobile app and only made their first online purchase after activation, so contamination with *organic adopters* (those adopted online shopping before the loyalty program was launched and then activated the loyalty program after it was launched) is unlikely.

2.8.2 Doubly robust DID, Propensity Score Matching and IPW approaches

In §2.5.3, we saw that there are some differences in the demographics and other pre-adoption variables between (1) adopters and *offline-only* customers, and (2) *organic adopters* and *event adopters*. Therefore, to provide additional credibility to our findings and ensure that they are robust to any factors correlated with observable characteristics that may affect the propensity to adopt online shopping organically versus spurred by external events, we also consider doubly robust DID (Sant'Anna and Zhao, 2020) and propensity score-based approaches (Rosenbaum and Rubin, 1985; Stuart et al., 2011). These methods provide better comparability between groups by improving covariate balance by discarding poorly matched customers. We compute propensity scores using a logistic regression model between (1) adopters vs. *offline-only* customers, and (2) *organic adopters* vs. *event adopters*, with variables including demographics (e.g., gender, age, and household income) and pre-adoption

[13]The ATT point estimates are taken from Tables 2.7 and 2.10. The standard errors are obtained from 500 bootstrap replications. The bootstrap distribution of "true" ATT of *Black Friday adopters* relative to *organic adopters* is presented in Figure A.7 in Appendix §A.7.1.

purchase variables (i.e., average monthly spend, number of unique orders, unique brands, and unique categories). With the estimated propensity scores, we apply our DID estimator with both IPW and PSM methods and also run a doubly robust DID estimator. The details of the analysis and the results are shown in Web Appendices §A.7.2 for *organic adopters* vs. *event adopters* and §A.7.3 for *adopters* vs. *offline-only* customers. We find that our substantive conclusions remain the same with these approaches.

2.8.3 Alternate definitions for the organic adopter and covid adopter cohorts

To check whether our results are sensitive to the particular months we chose to define our *covid adopter* and *organic adopter* cohorts in our COVID-19 analysis, and to test whether our conclusions would change if we used "later" *covid adopters* vs. "early" *covid adopters*, we repeat our analysis with alternative definitions for both cohorts and present results in Appendix §A.7.4. To start with, we redefine *organic adopters* as those customers who were offline-only before 2019-10-31 and made their first online purchase between 2019-11-01 and 2020-02-29 (i.e., we extend the period we previously used to define *organic adopters*). We re-run the same set of DID models and find that all results are consistent with our main analysis. Next, we redefine the cohort of *covid adopters* as those customers who were offline-only before 2019-12-31 and made their first online purchase between 2020-05-01 and 2020-06-30 (i.e., two months later than we had used in our main analysis). Again, we find that all the results are consistent with the main analysis. Finally, we compare "early" *covid adopters* (who adopted online shopping between 2020-03-16 and 2020-04-30) and "later" *covid adopters* (who adopted online shopping between 2020-05-01 and 2020-06-30). Again, we run the same set of DID models and do not find any significant differences except that the "later" *covid adopters* spend slightly higher by 2.48%.

These analyses also address another potential concern with our main analysis, which is that *organic adopters* adopt for two extra months relative to *covid adopters* before our common post-adoption period, and this may drive some of the results. The fact that our conclusions don't change when we change the set of months used to define our cohorts also allays this concern. Finally, our definitions for *Black Friday adopters* and *Loyalty Program adopters* are much more precise than for *covid adopters*, so we don't repeat this robustness check for those analyses.

2.8.4 Alternate operationalizations of dependent variables

To test whether the positive skewness in purchase outcomes affects our results, we re-estimate all DID models using a natural log transformation for the dependent variables ($\ln(y+1)$) and report the results in Appendix §A.7.5. Across all specifications, the results remain consistent with our main findings. Another potential concern in our setting is that many customer-month spending observations are zero in our data. Therefore, we consider an alternative specification that accommodates zero-inflated dependent variables. Specifically, we estimate a two-part (hurdle) model at the customer-month level (Cameron and Trivedi, 2005). This approach involves: (1) a linear probability model with a DID specification to estimate changes in the likelihood of making a purchase during the post-adoption periods[14]; (2) a DID model that estimates the changes in spend conditional on making a purchase.[15] Overall, we find that the results from these specifications are consistent with our main results; see Appendix §A.7.6 for detailed estimation results.

2.9 Conclusions

We study how the adoption of a retailer's online channels by previously *offline-only customers* affects post-adoption behavior, and whether these effects depend on the *reason* for adoption. Prior work shows that multichannel customers are more valuable than single-channel customers, but largely treats channel adoption as homogeneous. We show that this masks important heterogeneity. Using transaction data from a large Brazilian pet supplies

[14]Here, the dependent variable of linear probability model is $\mathbb{I}\{Spend_{it}\}$, a binary indicator for whether customer i made purchase during month t, and the DID specification is Equations (2.1) for comparing *adopters* vs. *offline-only* customers, and Equation (2.2) for comparing *organic adopters* vs. *event adopters*)

[15]We estimate the same equations as the main analysis, but use the customer-month observations with non-zero spend.

retailer from 2019–2024, we compare customers who adopt online shopping organically with those induced to adopt through three external pathways: COVID-19, Black Friday promotions, and a mobile-app-activated Loyalty Program.

Main findings. We document three main findings. First, online adoption is associated with higher spend and profitability relative to remaining offline-only, consistent with the standard multichannel-customer result. Second, conditional on adoption, post-adoption behavior differs sharply across adoption pathways. *Covid adopters* spend similarly to *organic adopters* but are more profitable, because they continue allocating more of their spending to the higher-margin offline channel. This pattern is consistent with consumer inertia and habit persistence, especially among older customers. Third, promotion-induced adopters behave differently: *Black Friday adopters* and *Loyalty Program adopters* spend less and are less profitable than *organic adopters* after adoption. Event-study evidence shows that these customers spend more around the adoption period but less afterward, consistent with forward buying and post-promotion demand distortion.

The channel-mix results further show that external adoption does not always translate into sustained online usage. *Covid adopters* move toward online shopping more slowly than *organic adopters*. *Black Friday adopters* initially retain a higher offline share, but move online faster over time, consistent with their younger profile. *Loyalty Program adopters* do not differ meaningfully from *organic adopters* in offline spending share, likely because recurring app-based activation reinforces online engagement. Thus, adoption pathways affect not only how much customers spend, but also where they spend and how profitable they are.

Managerial implications. The results imply that firms should not value all online adopters using a common multichannel benchmark. In our setting, ignoring adoption pathways substantially biases incremental CLV forecasts and promotion ROI calculations. In particular, pooling adopters or using *organic adopters* as proxies can substantially overstate the long-run value of promotion-induced adopters and make online-adoption promotions appear to break even much faster than they actually do. Firms should therefore estimate incremental CLV separately by adoption pathway, set pathway-specific caps on acquisition and promotional costs, and design online-adoption campaigns that account for post-adoption behavior. For promotion-induced adopters, this means limiting stockpiling and encouraging repeat engagement after the promotion. For shock-induced adopters, it means using habit-reinforcing interventions to convert a temporary online trial into a sustained multichannel behavior.

Limitations and future research. Our study has three limitations that suggest opportunities for future work. First, we study one macro shock and two firm-driven promotions. Future research could examine other adoption triggers, such as store closures, weather shocks, localized disruptions, or different promotion designs. Second, our setting is a single retailer in one category and country; studying other categories and retail environments would help assess the generalizability of the mechanisms we document. Third, our estimates are descriptive ATTs that capture the post-adoption value of customers who actually adopt through each pathway. This is the estimand most relevant for managerial forecasting and campaign evaluation, but future work could complement our analysis by separately estimating causal ATEs of online adoption using randomized encouragements, instruments, or other identification strategies.

Chapter 3

HOW DOES ONLINE SHOPPING AFFECT OFFLINE PRICE SENSITIVITY?

3.1 Introduction

Over the past two decades, online shopping has grown rapidly, driven by the widespread adoption of computers and smartphones. U.S. e-commerce sales more than doubled from $571 billion in 2019 to approximately $1.19 trillion in 2024, and the share of e-commerce in total US retail sales has grown from 15.4% to 22.7% over the same period (Haleem, 2025). Consumers are increasingly adapting to the convenience of online shopping and allocating a large portion of their spending to online channels (McKinsey & Company, 2023). This shift has prompted traditional brick-and-mortar retailers such as Walmart, Target, and Costco to increase investments into their e-commerce channels as a major driver of growth.

Given the rise of online shopping and investments in e-commerce, a natural question that arises for retailers is whether a consumer's adoption of a given retailer's online channels influences her shopping behavior at that retailer's physical stores, and if so, how? In this paper, we focus on one important aspect of consumer behavior – their price sensitivity. Specifically, we seek to answer the following research question: When an offline-only consumer starts using the online shopping channel(s) of a given retailer, does it affect her offline price sensitivity (i.e., price sensitivity at that retailer's physical stores)? If so, does offline price sensitivity increase or decrease post-online adoption? There are two main reasons why answering this question is important from both researchers' and managers' perspectives. First, from a substantive perspective, understanding the determinants of price elasticity is important – doing so gives us insights into what factors make consumers more or less price sensitive, and how. Second, from a managerial perspective, quantifying the spillover effects of online shopping on offline price sensitivity can have significant implications for retailers' pricing strategies and profits. Given the migration of consumers from offline stores to multi-channel shopping, quantifying the existence and impact of cross-channel externalities can help firms optimize and coordinate pricing and promotion across channels. More broadly, there is a growing interest among managers in understanding how interactions and spillovers between online and offline shopping channels shape consumer behavior, given the growth of omnichannel retailing across the globe (Zhang et al., 2019; Biswas et al., 2024).

While there is a rich literature studying the determinants of price sensitivity, this literature largely focuses on determinants of price elasticities *within* a channel (see §3.2 for a detailed discussion). In contrast, our focus is on the externalities or spillovers *across* channels. From a conceptual perspective, there are reasons why we may expect offline price sensitivity to increase post-online adoption as well as reasons that suggest the opposite. On the one hand, online shopping may make consumers less price sensitive and more brand loyal because online recommendation systems promote easy repeat purchases and lead to habit formation by highlighting prior purchases, e.g., "Buy again" or "Frequently purchased items" (Danaher et al., 2003). This increase in brand loyalty in a retailer's online channels could spill over to offline shopping occasions and may make consumers more brand loyal and less price sensitive in that retailer's physical stores too (after adopting online shopping). On the other hand, online shopping makes it easy for consumers to engage in price comparisons (due to features like sorting by price and lower search costs of browsing through many options on an app or website). This habit of checking and comparing prices on a retailer's online channels can potentially carry over to the retailer's offline channels, and consumers may become more aware of price differences in the retailer's physical stores as well. Another potential reason that may make consumers more price sensitive after adopting online shopping is that, in many contexts, prices are lower online than offline (Brynjolfsson and Smith, 2000; Cooper, 2006). Being exposed to lower prices for the same products online can make in-store prices seem inflated by comparison, prompting greater scrutiny of the prices of different brands. Finally, recent research suggests that when consumers are exposed to a wider variety of prices (which is likely to occur when they adopt online shopping), they become more price sensitive (Aparicio et al., 2024). Therefore, the question of whether the adoption of a retailer's online channels affects consumers' offline price sensitivities at that retailer's physical stores, and if so, in which direction, is an empirical one.

We address this question using data from a large pet supplies retailer in Brazil that sells through both online and offline channels. The dataset includes detailed transaction-level information for all customers across both channels from January 1, 2019, to June 30, 2021. In our data, we observe some consumers who were initially offline-only and then made their first online purchase with the retailer within our data window, while others remained offline-only throughout. This allows us to track the offline purchases of online shopping adopters, both pre- and post-adoption, and also all purchases of those consumers who remained offline only. We focus on four major product categories: dog food, cat food, dog hygiene, and cat hygiene, which together account for more than 50% of the retailer's total revenue. These categories are also relatively frequently purchased and are characterized by a good amount of price variation both across brands and across time, which allows us to measure individual-level price elasticities at the monthly level.

From an empirical perspective, a key question relates to the estimand of interest and how to measure it. Note that adopting online shopping is a consumer-level decision, i.e., firms cannot randomly assign consumers to adopt online shopping. Therefore, we cannot quantify the Average Treatment Effect (ATE) of adopting online shopping on offline price sensitivity. In addition, even if obtaining the ATE was possible, firms cannot act on it because they cannot force consumers to adopt online shopping. Therefore, obtaining the ATE is not only impractical but also lacks actionable implications. Hence, we focus on identifying the Average Treatment Effect on the Treated (ATT), which captures the impact of treatment on those who adopted online shopping.

Our empirical strategy consists of two steps. In the first stage, we specify and estimate a discrete choice logit model for each category, controlling for unobserved customer heterogeneity in brand preferences. In our model, we allow consumers' offline price sensitivity to be a function of the extent to which a given consumer adopts online shopping at a point in time, measured by the cumulative fraction of their online spending with the retailer up to that point in time.[1] This modeling approach allows us to measure the effect of interest and examine how the price elasticity of consumers who have adopted online shopping diverges over time from that of those who have not. We encounter two challenges in estimating this demand model. First, maximum likelihood estimation of a non-linear model, such as the logit model, is quite challenging in the presence of high-dimensional fixed effects (customer-brand fixed effects in our case). To accommodate these high-dimensional fixed effects, we build on the Minorization-Maximization algorithm proposed by Chen et al. (2022) and allow for covariates that vary by individual, by alternative, and also over time. We then apply this algorithm to our setting for estimation. To the best of our knowledge, this is the first empirical paper in marketing to efficiently estimate high-dimensional customer-brand fixed effects in a demand model using the MM algorithm. Second, we must address potential endogeneity in prices using our observational data (Nevo, 2001). To that end, we use costs as an instrument (Berry, 1994) in conjunction with a control function approach (Villas-Boas and Winer, 1999; Petrin and Train, 2010). We then use the estimates from this model to compute customer-month-level price elasticities.

In the second stage, we use the estimated customer-month-level offline price elasticities of adopters and non-adopters from the first stage to estimate our ATT of interest, which is the effect of adopting online shopping on consumers' offline price elasticities. Again, our observational data pose some challenges for the identification of the ATT. First, consumers in our sample are either treated or untreated in each time period, and we do not observe their counterfactual outcomes in the unobserved condition (i.e., their behaviors if they had not adopted online shopping). Moreover, the staggered timing of online adoption among different consumers requires careful computation and aggregation of treatment effects (Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Baker et al., 2022). To deal with this, we

[1] We also include an alternate specification with a binary online adoption variable, as a robustness check.

employ a staggered difference-in-differences (DID) framework, as proposed by Callaway and Sant'Anna (2021), utilizing the variation in staggered adoption timing of online shopping. We identify different cohorts of consumers who adopt online shopping (i.e., adopters) in each month of our data and compare changes in their behavior to those of a control group of customers who shop exclusively offline throughout the observation period (i.e., non-adopters).

Our main finding is that consumers become *more* price sensitive in their offline shopping trips after adopting online shopping for three out of four categories. The relative increase in offline price elasticity, defined as the ATT relative to the pre-adoption average price elasticity of the treated group is heterogeneous across categories. We find a large effect in the dog and cat hygiene categories (where price elasticity increases by 72.7% and 34.7%, respectively). We also observe a significant but smaller increase in the dog food category (8.4%) and no effect for cat food. The heterogeneity across categories can likely be explained by differences in switching costs to different brands. While pet owners face little risk in trying a new lower-priced hygienic mat for their pet, changing the brand of food can be much trickier due to the pets' stronger preferences for shape, smell, and texture in food, particularly for cats (Nichelason, 2024).

We also provide a series of tests and alternative specifications that establish the robustness of our results and provide additional empirical support for the generalizability of our findings. First, we show that using a binary online adoption variable in our demand model does not change our conclusions. Second, we also use Inverse Propensity Weighting (IPW) approaches to control for self-selection of the online adoption decision. Third, we consider an alternative estimator where later adopter groups serve as a control group for earlier adopters. Fourth, we consider an alternate definition of purchase occasions to also include online shopping visits. Our conclusions and substantive findings remain unchanged across all of these robustness checks. More broadly, we also present some evidence in favor of portability to other settings. For instance, our data period also includes COVID-19, an exogenous macro-environmental shock. Interestingly, we do not find differences in the adoption effect between customers who adopted online shopping during COVID-19 vs. those who adopted organically before that, indicating that the effect is not sensitive to the reasons/drivers of a consumer's decision to adopt online shopping.

Our paper contributes to the broad literature on online-offline retailing and pricing. From a substantive perspective, we contribute to the literature on price sensitivity and multichannel retailing by identifying a new determinant of offline price sensitivity: the adoption of online shopping. We demonstrate that consumers become more price sensitive in a retailer's physical stores after adopting that retailer's online channels for shopping. This finding is important and distinct from previous research, which has mainly focused on cross-sectional or within-household comparisons of price sensitivity between online and offline settings. From a managerial perspective, our findings have implications for firms' pricing strategies and profitability, emphasizing the need for multichannel retailers to account for the online adoption effect (i.e., heightened offline price sensitivity when consumers adopt a retailer's online channels) in their pricing decisions.

The rest of the paper is organized as follows. In §3.2, we discuss the related literature, and in §3.3, we discuss the setting and data. In §3.4, we document descriptive evidence, and in §3.5, we specify the firm's problem and provide an overview of the empirical strategy. §3.6 presents the demand model, and §3.7 presents the ATT estimates of online adoption on offline price sensitivity. §3.8 discusses the generalizability and robustness checks of our main results. Finally, in §3.9, we summarize our findings and their implications, and also discuss avenues for future research.

3.2 Related Literature

Our study relates to multiple streams of literature on the determinants of price elasticity, digital platforms, and multi-channel retailing.

First, our work relates to the large literature in marketing and economics that has focused on documenting the drivers of consumers' price sensitivity. Prior research has identified various determinants of price sensitivity, particularly in grocery settings. These include (1) macroeconomic conditions, such as GDP growth and inflation rates (Gordon et al., 2013; Tellis, 1988; Bijmolt et al., 2005); (2) marketing interventions, such as advertising (Kanetkar et al., 1992; Kaul and Wittink, 1995; Kalra and Goodstein, 1998; Jedidi et al., 1999), promotions Mela et al. (1997); Jedidi et al. (1999); Kopalle et al. (1999) and category expansion (Yu et al., 2021); (3) brand strength (Guadagni and Little, 1983; Kamakura and Russell, 1989; Guadagni and Little, 2008); (4) consumer loyalty (Danaher et al., 2003); (5) time between purchase and consumption (Joo et al., 2020). Early research has also examined the relationship between customer demographic characteristics – age, education, income, and household size – and price sensitivity using household-level scanner panel data (Rossi and Allenby, 1993; Narasimhan, 1984; Bawa and Shoemaker, 1987, 1989). However, later studies have concluded these relationships are generally weak, suggesting that demographics do not explain a large portion of the variation in household-level price sensitivity (Hoch et al., 1995; Rossi et al., 1996).

We consider the adoption of a given retailer's online channels as a novel factor that may affect consumers' offline price sensitivity at that retailer's physical stores. Given the rise of digitization and online shopping, this factor can be particularly important for multi-channel retailers (who are increasingly investing in online channels). Existing literature provides some insights into how consumer behavior differs between online and offline channels. Early studies focusing on cross-sectional comparisons find that consumers exhibit lower price sensitivity and less brand switching when shopping online compared to offline (Degeratu et al., 2000; Andrews and Currim, 2004). Danaher et al. (2003) shows that high market share brands exhibit greater brand loyalty online, whereas low market share brands exhibit greater brand loyalty in offline stores. However, these studies are "cross-sectional" in nature, i.e., they compare the set of consumers shopping offline vs. those shopping online, who may be very different from each other. More recent research has shifted toward within-household analysis. For example, Chu et al. (2008) compare price sensitivity in online and offline grocery shopping within the same household and find that consumers are generally less price sensitive online. Further, Chu et al. (2010) investigate the role of household characteristics in moderating price sensitivity across channels. However, none of this research looks into *if* and *how* the *adoption* of online shopping affects consumers' offline price sensitivity, which is the focus of our paper.

More broadly, our study contributes to the broader multichannel retailing literature; see Cui et al. (2021) for a comprehensive overview. One key stream of research highlights that online and offline channels can function as complements (Avery et al., 2012; Wang and Goldfarb, 2017; Bell et al., 2018) or substitutes (Forman et al., 2009; Li, 2021; Shriver and Bollinger, 2022), depending on the context. Other studies have examined how digitization affects consumer shopping behavior across channels, focusing on spend patterns and purchase volume (Wang et al., 2015; Narang and Shankar, 2019; Biswas et al., 2024), brand exploration (Pozzi, 2012), and purchase variety (Chintala et al., 2023). Building on this body of work, our study provides new insights into multichannel strategy (Neslin, 2022) by investigating changes in offline price sensitivity when consumers adopt a retailer's online channels for shopping.

3.3 Setting and Data

3.3.1 Setting

We use data from a leading pet supplies retailer in Brazil. This firm operates in a multichannel environment, with over 200 brick-and-mortar stores nationwide and a digital presence through both a mobile app and an e-commerce website. The retailer offers a broad selection of pet-related products, including food, medicines, hygiene and grooming items, toys, and accessories. To track consumer purchases across online and offline channels, the retailer uses Brazil's National Identification Number, which provides a unique identifier for each customer. Online purchases - whether through the app or website - require customers to log into their accounts, making all such transactions traceable to individual IDs. For brick-and-mortar purchases, customers provide their ID at checkout, although it is not mandatory. According to the retailer, around a quarter of offline transactions are not associated with a customer

ID, and this rate has remained stable throughout the study period.[2]

The dataset spans 30 months, from January 1, 2019, to June 30, 2021, and covers all transactions (online and offline) linked to identifiable customers. Each transaction includes a unique order ID and detailed product-level information, such as price, quantity, brand, category (e.g., dry food or hygiene), and purchase channel.

3.3.2 Dataset Construction

We now discuss how we construct the dataset used for our empirical analysis.

Selection of Product Categories: We focus on the four top-selling non-medical categories: Dry Dog Food, Dry Cat Food, Dog Hygiene (i.e., Hygienic Mats), and Cat Hygiene (i.e., Sands). These categories account for more than 50% of the focal retailer's revenue (see Table 3.1). Products within these four major categories are generally measured in comparable units, allowing us to construct brand-level prices per unit.

We exclude certain categories from our analysis. First, we exclude medical products such as prescription drugs and anti-flea medication (second category shown in Table 3.1) because many products in this category require a doctor's prescription and purchase does not depend solely on consumers' preferences. Further, medical products can be quite differentiated and target specific ailments, and as such, do not serve as natural substitutes for each other. Second, we exclude all small categories that account for less than 3% of the total sales revenue during the observation period, e.g., snacks and toys. Products in these categories are often unique, serve diverse functions, and are not necessarily natural substitutes for each other.

Selection of Brands and Stock Keeping Units (SKUs): Each category contains a large number of pack sizes and SKUs per brand. First, within each category, we focus on specific pack size(s) to ensure that the price variation for each brand is captured accurately – this avoids the need to aggregate price movements across a number of different pack sizes.

[2]The fact that this rate has remained stable despite the fact that there was a spike in online shopping adoption during COVID-19 reassures us that the adoption of online shopping does not affect it.

Table 3.1. Top Categories Ranked By Total Sales (Offline and Online) During the Observation Period.

Category	Share of sales(%)	Cumulative share (%)
Dry Dog Food	32.62	32.62
Medical	12.18	44.80
Dry Cat Food	9.41	54.21
Dog Hygiene	5.48	59.68
Cat Hygiene	3.62	63.30

Focusing on a narrow set of pack sizes to study brand choice within a category is common in the brand choice literature (e.g. Allenby and Rossi (1998); Kim et al. (2002); Chintagunta et al. (2005); Dubé et al. (2010)). The pack sizes we select for each category are as follows: 15KG for dry dog food, 10 KG for dry cat food, 80cm $\times$ 60cm 30-piece mats for dog hygiene, and silica cat sand (1.6KG, 1.8KG, and 2.0KG) for cat hygiene. We provide a detailed discussion of the reasons for selecting these pack sizes in Appendix §B.1.1.1.

Next, we aggregate SKUs to the brand level and estimate consumer preferences and price elasticities at the brand level for each product category, similar to Hitsch et al. (2021); Gordon et al. (2013). Specifically, we follow a two-step process for each category. First, for each category, we select the top-selling brands and SKUs that are comparable in usage and are substitutes for each other. Next, we aggregate prices and costs (our instrumental variable to deal with price endogeneity) to the brand level from the SKU level, following the approach described in Hitsch et al. (2021). We provide details on both these steps in Appendix §B.1.1.

Selection of Customers: Following the standard approach in the brand choice literature (Kim et al., 2002; Chintagunta et al., 2005; Dubé et al., 2010), for each category, we restrict our sample to only those consumers who only purchase the specific pack size of interest for that category (as defined earlier), without switching to any other pack sizes within the same category. This is so that we are able to fully capture these customers' substitution patterns within the category. Next, given that our goal is to investigate the effect of consumers' adoption of online shopping at a given retailer[3] on their *offline* price sensitivity at that retailer's physical stores, we need to ensure that we have sufficiently long pre- and post-adoption periods for each adopter in our sample to effectively measure price elasticities pre- and post-adoption. Therefore, we focus on customers who adopt online shopping with the retailer in the middle of our observation period, i.e., July 1st, 2019 – June 30th, 2020. So our 30-month data period is divided into three phases: a pre-adoption period (January 1st, 2019 – June 30th, 2019) when all adopters are in their pre-adoption phase, an adoption period (July 1st, 2019 – June 30th, 2020) during which time the adopters in our sample make their first online purchase, and a post-adoption period (July 1st, 2020 - June 30th, 2021), when all the adopters in our sample are in their post-adoption phase. To clarify, a given adopter's pre- and post-adoption phases are defined as the months before and after they make their first online purchase. Non-adopters only shop offline through the entire data period.[4]

We now outline the criteria used to select adopters and non-adopters for each product category below. *Adopters* in a given category must satisfy the following criteria:

- Made their first-ever online purchase (in any product category, not just the focal category) between July 1st, 2019, and June 30th, 2020.
- Made at least one offline purchase in the focal category both before and after online adoption.

Overall, this gives us 12 cohorts of adopters, where each cohort represents the month in which the user adopted online shopping. For example, the first cohort of adopters consists of users who adopted online shopping during July 2019 while the twelfth cohort consists of users who adopted online shopping during June 2020. Figure B.1 in Appendix B.1.1.3 shows stylistic examples of users in different cohorts. Figure 3.1 shows the number of adopters across the different periods in each category. Notice that there is a spike in the number

[3] "Adoption of online shopping" refers to a customer making his/her first online purchase with the focal retailer in *any* category, and not just one of the four focal categories.

[4] Note that we also verified with the retailer that the identified adopters and non-adopters in our data did not adopt online shopping before our data period begins.

of adopters during March–April 2020, which is likely driven by the onset of the COVID-19 pandemic. Our analysis is agnostic to the reasons for adopting online shopping, and in our empirical analysis, we will allow for the possibility that these cohorts may be different from earlier/later cohorts.

Non-adopters in a given category must meet the following criteria:

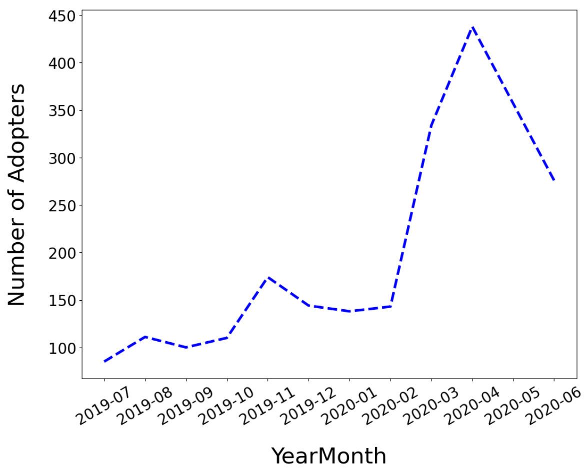
- Did not make any online purchases at any point during the entire observation period.
- Made at least one offline purchase in the category during both the pre-adoption (i.e., January 1st, 2019 – June 30th, 2019) and the post-adoption periods (i.e., July 1st, 2020 – June 30th, 2021).

Estimation Sample Construction: For each customer in a given category (both adopters and non-adopters), we organize the data into a panel that captures purchase occasions and brand choices for individual customers over time. Specifically, we define purchase occasions at the customer-month level. If a customer visits a physical store of the retailer during a month and purchases in any category (not limited to the categories we focus on in this study), that customer-month is included as a purchase occasion for all four categories. If a customer makes no purchases in any category during the month, that customer-month is not included as a purchase occasion. In other words, a purchase occasion represents any customer-month during which we observe any offline interaction with the retailer (similar to week-store visit in Gordon et al. (2013)).[5] This approach of using purchase occasions to construct the estimation sample is standard in the demand estimation literature studying brand choice using household level data, since it allows researchers to focus on modeling brand choice and purchase incidence in a category conditional on a store visit, instead of modeling the decision of when to go to the store or which store to visit (Chib et al., 2004;

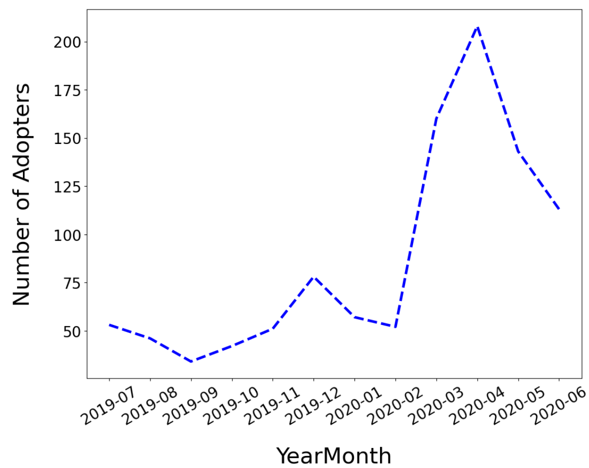
[5] Based on our definition of purchase occasion, if a customer visits multiple times in a month, it is treated as one purchase occasion. In our data, customers, by and large, do not purchase from more than one store in a month. More than 96% of customer-month-level purchase occasions involve purchases from a single store. When we process the choice data for each category, for the minor fraction of purchase occasions that involve more than one store, we label these customer-month purchase occasions with the store where the customer spent the most within that category during the month so that each customer-month level purchase occasion is associated with a single store.

Figure 3.1. Sample Sizes of Adopters by Cohorts and Product Category. Corresponding sample sizes for each cohort are shown in Table B.1 in Appendix B.1.1.3.

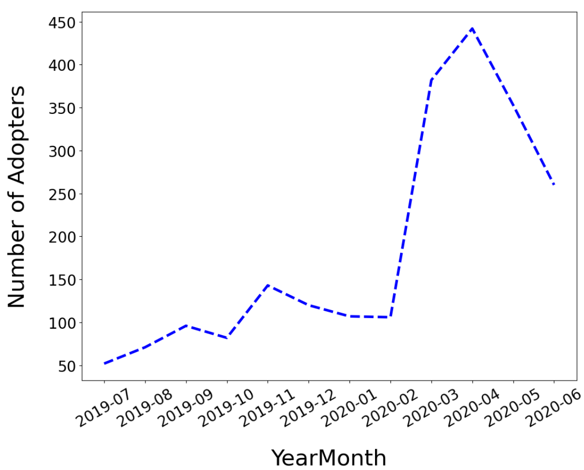
(a) Dry Dog Food



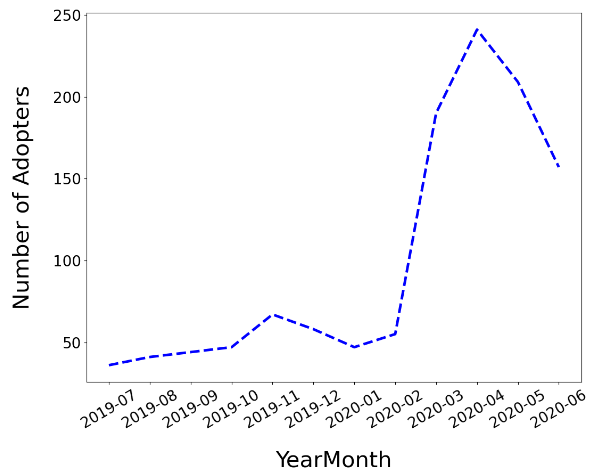
(b) Dry Cat Food



(c) Dog Hygiene



(d) Cat Hygiene



Chintagunta et al., 2005; Dubé et al., 2010; Gordon et al., 2013). Following the standard practice in the literature, we do not model the channel choice decision explicitly for a couple of reasons – (a) channel choice can depend on multiple factors that we do not observe (e.g. transportation costs, distance to stores, consumer expectations of prices etc.); (b) we would need to impose a structure on the order of store choice, channel choice and brand choice and it is ex-ante not clear in which order consumers make these decisions. Given that our focus is on documenting the empirical effect of the adoption of online shopping on price sensitivity in offline brand choice decisions, modeling the arrival of purchase occasions is neither central nor necessary to answer our research question. Further, in §3.8.2.4, we show that our results are robust to alternative definitions of purchase occasions.

For each purchase occasion, we identify the brand chosen by the consumer as the brand that has positive spend.[6] If the customer does not have positive spend on any brand during a purchase occasion, we label it as choosing the outside option (i.e., "no-purchase").

3.3.3 Summary Statistics

We now present some summary statistics on the estimation sample.

First, we find that there are systematic differences in the demographics (age, income, and gender) of adopters and non-adopters across all four categories. Across all categories, women tend to adopt online shopping more than men. Additionally, online adopters tend to be younger and from higher-income households. This makes intuitive sense because younger and higher-income consumers are often more familiar with online technology and seek time-saving and convenient shopping options. Details of these demographic characteristics are

---

[6]We rarely observe customer-month purchase occasions that involve a purchase of more than one brand. All product categories have fewer than 0.86% of total purchase occasions (e.g., including those with multiple visits to a single store in a month, visits to multiple stores and a single visit) with multiple brand purchases. For these few purchase occasions with multiple brand purchases, we assign the brand with the largest spend for that customer during the month, similar to previous brand choice literature, e.g. Erdem and Keane (1996). Some researchers have handled this issue differently, for instance, by having each brand chosen count as an independent single-brand purchase occasion (Krishnamurthi and Raj, 1988; Chib et al., 2004). Since the fraction of purchases is so small, we do not think this decision makes a material impact on our results.

shown in Appendix §B.1.2.

Next, Table 3.2 summarizes the category-level statistics, including the number of brands, sample sizes for adopters and non-adopters (as defined above), the total number of purchase occasions, and the number of purchase occasions that contain a brand purchase within each category.

Table 3.2. Summary Statistics By Category

	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Number of Brands	8	4	4	4
Number of Adopters	2,410	1,037	2,214	1,192
Number of Non-Adopters	7,938	2,297	3,582	2,725
Number of Purchase Occasions	136,411	45,443	78,253	55,410
Purchase Occasions with a Category Purchase	99,233	27,623	40,183	42,596

Table 3.3. Summary Statistics of Prices and Costs

	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Mean Price	10.17	11.89	2.28	12.25
SD Price	2.82	2.40	0.28	2.46
Mean Cost	6.29	8.04	0.96	4.95
SD Cost	2.32	3.39	0.29	1.54

Finally, in Table 3.3, we show some summary statistics of the prices and costs for each category. Specifically, we show the mean and standard deviation (SD) of the brand price (or cost) per pack size unit (e.g, KG or piece) for each category, for the estimation sample. These values are calculated by averaging across all purchase occasions and brands. Specifically, the mean price for category c is calculated as: $\text{Mean Price}_c = \frac{1}{N_c}\frac{1}{J_c}\sum_{t=1}^{N_c}\sum_{j=1}^{J}\text{Price}_{jt}$ where Price_{jt} is the price of brand j on the purchase occasion t. N_c is the total number of purchase occasions in category c, and J_c is the number of brands of category c. The standard deviation is computed as $\text{SD Price}_c = \sqrt{\frac{1}{N_c}\frac{1}{J_c}\sum_{t=1}^{N_c}\sum_{j=1}^{J}(\text{Price}_{jt} - \text{Mean Price}_c)^2}$. The mean cost and SD cost are defined similarly.[7] Overall, we see that the data contain significant price and

[7]All price and cost figures are inflation-adjusted. Brazil exhibited significant inflation during our study period, and adjusting for inflation is crucial to reflect the real price changes. The adjustment is relative

cost variation for each category. This variation is important and helps us estimate price elasticities.

3.4 Descriptive Analysis

We now present some descriptive analysis that examines if/how the adoption of online shopping with the retailer changes consumer behavior during their offline shopping trips at the retailer's physical stores in the post-adoption period. Specifically, we focus on two variables that can capture changes in *within-customer* offline price sensitivity:

- Price Paid: Conditional on making an offline purchase in the focal category, do adopters shift toward lower-priced after online adoption (compared to non-adopters)?
- Category Purchase Frequency: Conditional on an offline purchase occasion (as defined earlier), do adopters purchase the focal category less frequently after they adopt online shopping, compared to non-adopters?

3.4.1 Price Paid

We examine how the prices paid by adopters change after online adoption (compared to non-adopters). To that end, we focus on all offline purchase orders (made by adopters and non-adopters) that contain a focal category purchase and estimate the following two-way fixed-effects regression for each category at the customer-order level:

$$\ln(\text{Price\_Paid})_{it} = \delta \times \mathcal{I}\{\text{Adopter}_i\} \times \mathcal{I}\{t \geq T_i^{\text{adopt}}\} + \text{Customer}_i + \text{YearMonth}_t + \epsilon_{it}, \quad (3.1)$$

where $\mathcal{I}\{\text{Adopter}_i\}$ is an indicator variable which is 1 for adopters and 0 for non-adopters. T_i^{adopt} is the adoption time of customer i, and $\mathcal{I}\{t \geq T_i^{\text{adopt}}\}$ is an indicator which is 1 if the order t is during customer i's post-adoption period. We also include customer-fixed effects, Customer_i, to measure the within-customer change, and time-fixed effects, YearMonth_t, to control for time-specific unobserved shocks. A negative δ indicates that adopters, on average,

to the first observation period, January 2019.

Table 3.4. Two-Way Fixed Effects Model Estimates – ln(Price_Paid)

Dependent Variable:	ln(Price_Paid)			
	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)	(4)
Variables				
$\mathcal{I}\{\text{Adopter}_i\} \times \mathcal{I}\{t \geq T_i^{\text{adopt}}\}$	-0.0102*	-0.0106**	-0.0046	-0.0160*
	(0.0040)	(0.0038)	(0.0030)	(0.0063)
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	99,233	27,623	40,183	42,596
R^2	0.89566	0.70460	0.62699	0.66927
Within R^2	0.00030	0.00065	0.00012	0.00074

Clustered (Customer & YearMonth) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

pay lower prices or shift towards the lower end of the price distribution after adopting online shopping, relative to non-adopters. A positive δ would indicate the opposite.

Table 3.4 presents the within-customer changes in the prices paid by adopters relative to non-adopters, controlling for customer and date fixed-effects. Across all product categories, we see that adopters pay lower prices offline after they adopt shopping online in the retailer, compared to non-adopters. Further, three of the categories show a statistically significant effect at the 5% level of significance. In Appendix B.2 Table B.3, we repeat this analysis with the percentile of price paid as the dependent variable (instead of ln(Price_Paid)) and find consistent results – customers shift towards lower-priced products during their offline purchases after adopting online shopping. Together, these results suggest that, conditional on purchasing, consumers are paying lower prices, possibly because they are more price sensitive.

Table 3.5. Two-Way Fixed Effects Model Estimates - BuyCategory

Dependent Variable:	BuyCategory			
	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)	(4)
Variables				
$\mathcal{I}\{\text{Adopter}_i\} \times \mathcal{I}\{t \geq T_i^{\text{adopt}}\}$	-0.0751***	-0.0497***	-0.0617***	-0.0654***
	(0.0077)	(0.0109)	(0.0075)	(0.0106)
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	136,411	45,443	78,253	55,410
R^2	0.30479	0.25032	0.23015	0.26662
Within R^2	0.00147	0.00066	0.00113	0.00121

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

3.4.2 Category Purchase Frequency

Next, we examine how the frequency of adopters' purchase of a given category in the offline channel changes compares to non-adopters. Specifically, during each offline purchase occasion, we construct a binary dependent variable, $\mathcal{I}\{\text{BuyCategory}_{it}\}$, which equals 1 if customer i buys the focal category in purchase occasion t, and 0 otherwise. We then estimate the following two-way fixed effects regression at the customer-month purchase occasion level:

$$\mathcal{I}\{\text{BuyCategory}_{it}\} = \delta \times \mathcal{I}\{\text{Adopter}_i\} \times \mathcal{I}\{t \geq T_i^{\text{adopt}}\} + \text{Customer}_i + \text{YearMonth}_t + \epsilon_{it} \tag{3.2}$$

As before, we include customer and year-month fixed effects. The coefficient δ captures the change in the probability of buying the category for adopters relative to non-adopters after adoption.

Table 3.5 shows the results on this estimation. We find that, conditional on a purchase occasion, adopters purchase dog food and cat food 7.5% and 5.0% less frequently than non-adopters in the offline channel after they begin using the retailer's online channel. Similarly, adopters buy dog and cat hygiene 6.2% and 6.5% less frequently than non-adopters, respectively. These findings suggest that, compared to non-adopters, adopters are more likely to choose the "no-purchase" option when they start shopping online.

In summary, both the descriptive analyses suggests that consumers are less likely to buy the focal category offline and when they do buy, they pay lower prices, on average. Together these findings provide some suggestive evidence for potentially higher offline price sensitivity after online adoption. In the next few sections, we build on these findings and estimate customer-level price elasticities as a function of their online adoption status and then quantify the effect of online adoption on offline price elasticities.

3.5 Firm's Problem and Overview of Empirical Strategy

The primary objective of the firm is to quantify the causal effect of consumers' adoption of its online channels on their offline price sensitivity at its physical stores, compared to consumers who shop exclusively offline. However, quantifying this impact presents both methodological challenges and limited managerial relevance for several reasons.

First, adopting the online channel is a consumer-level decision, and firms cannot randomly assign consumers to adopt the online channel or remain offline-only through field experiments. This makes it impossible to directly measure the Average Treatment Effect (ATE) of online adoption on consumers' offline price sensitivity. Second, although firms could implement an intent-to-treat (ITT) design - such as randomly offering incentives to some offline-only consumers to encourage their online adoption - this approach still does not give us ATE given that some consumers would never adopt online [8].Third, even if a firm could somehow identify a significant effect of adopting online shopping on consumers' offline

[8] We do not obtain ATE under the ITT design, since the causal quantity obtained from 2SLS estimators represents the Average Causal Response, which is a weighted sum of individual treatment effects where weights are compliance scores; For more details, see Angrist and Imbens (1995); Syrgkanis et al. (2019); Chen et al. (2024); Mummalaneni et al. (2025).

price sensitivity, it cannot force consumers to choose a specific channel. Therefore, obtaining the ATE is not only impractical but also lacks actionable implications. Instead, the main quantity of interest for the firm is the change in offline price sensitivity among consumers who adopt online shopping compared to those who continue to shop exclusively offline i.e. the Average Treatment Effect on the Treated (ATT). Measuring this would help firms adjust their offline pricing strategies in response to their customers adopting online shopping. Thus, our goal is to identify the ATT of online adoption on offline price sensitivity.

We adopt a two-step empirical strategy to solve this problem:

- First, we measure category-level own-price elasticities at the customer-month level: In §3.6, we use a logit brand choice model where each consumer's price sensitivity coefficient varies with the extent of their online spending and include customer-brand fixed effects. To address price endogeneity, we use the control function approach (Petrin and Train, 2010), using wholesale cost as an instrument. To efficiently estimate the model given a large number of customer-brand fixed effects, we implement a Minorization-Maximization (MM) algorithm for the logit model (Chen et al., 2022).
- Next, we use the estimated elasticities in the first step to measure the change in offline price elasticities due to online adoption. In §3.7, we use a staggered difference-in-differences approach (Callaway and Sant'Anna, 2021) to estimate the ATT of online adoption on offline price elasticities after accounting for differences in the timing of adoption between different adopter cohorts.

3.6 Demand Model

3.6.1 Choice Model and Utility Specification

For each category, we use a logit model of brand choice (Guadagni and Little, 1983; Chintagunta et al., 1991) to estimate offline price elasticities at customer-brand-month-level.[9]

[9] As discussed in §3.3.2, we follow the standard approach in the brand choice literature and focus on the brand choice decision conditional on purchase occasion, and do not model the channel choice decision or the arrival of purchase occasions. Nevertheless, we note the results are robust to alternative definitions of

Suppose that for a given category, there are J brands indexed by $j = 1, 2, \ldots, J$ plus an outside option indexed by 0. The utility that consumer i obtains from choosing brand j during offline purchase occasion t is given by [10]:

$$\begin{aligned} U_{ijt} = {} & \alpha_0 \text{Price}_{jt} + \alpha_1 \text{cumPercOnlineSpend}_{it} + \alpha_2 \text{Price}_{jt} \times \text{cumPercOnlineSpend}_{it} \\ & + \alpha_3 \mathcal{I}\{\text{BuyPrevOcca}_{ijt}\} + \delta_{ij} + \epsilon_{ijt} \quad \forall j \in \{1, \ldots, J\} \end{aligned} \tag{3.3}$$

In this utility specification, Price_{jt} is the price of brand j during purchase occasion t. We allow a consumer's price sensitivity coefficient to depend on their cumulative percentage of online spending, denoted as $\text{cumPercOnlineSpend}_{it}$. This variable represents the proportion of the customer's total spending with the retailer that has occurred online up to the purchase occasion t since their first transaction with the retailer. We use this continuous variable to allow for variation in the online adoption effect based on the *extent* of a consumer's online spend. We also consider an alternate specification where we use a binary online adoption indicator in §3.8.2.1. The coefficient α_0 represents the baseline price sensitivity for purely offline customers. The interaction coefficient α_2 reflects how price sensitivity changes as consumers spend more online. A negative α_2 would suggest that higher online spending share leads to greater offline price sensitivity.[11] The variable $\mathcal{I}\{\text{BuyPrevOcca}_{ijt}\}$ is an indicator that equals 1 if customer i purchased brand j at the previous purchase occasion $t-1$, and 0 otherwise, capturing state dependence via α_3. The customer-brand fixed effects δ_{ij} capture consumer i's time-invariant preference for brand j. These fixed effects capture systematic time-invariant consumer-level heterogeneity in brand preferences. However, including these high-dimensional fixed effects introduces some estimation challenges which we discuss in §3.6.3. The term ϵ_{ijt} is an idiosyncratic error term. Finally, the utility of choosing the outside option is normalized as $U_{i0t} = 0 + \epsilon_{i0t}$ where ϵ_{i0t} also follows an i.i.d. Type I extreme value distribution.

---

purchase occasion; see §3.8.2.4 and Appendix §B.5.6.

[10] Here, we use a single index t to denote a purchase occasion, which contains both the time (YearMonth) and store information associated with the purchase occasion.

[11] Interaction terms in logit and other nonlinear models are best interpreted via marginal effects or elasticities (Ai and Norton, 2003).

3.6.2 Price Endogeneity

A common challenge in estimating price elasticities with observational choice data is price endogeneity. This issue arises when the error term, ϵ_{ijt}, contains some time-varying unobserved demand shocks that may be correlated with prices; e.g., if promotions or advertising are correlated with prices. To address this, we apply the control function approach for two endogenous variables, Price_{jt} and $\text{Price}_{jt} \times \text{cumPercOnlineSpend}_{it}$ (Petrin and Train, 2010; Wooldridge, 2015). We use the weighted wholesale cost indices for each brand j (constructed as described in §3.3.2), Cost_{jt}, as the instrument. Intuitively, cost is correlated with a brand's price but not correlated with unobserved demand shocks, and it is a standard instrument for correcting price endogeneity in the demand estimation literature (Berry et al., 1995; Berry and Haile, 2021).

To obtain control functions for two endogenous variables, we run the following two regressions and obtain the residuals; see (Aparicio et al., 2024) for a similar use case of this approach.[12]

$$\text{Price}_{jt} = \gamma_1 \text{Cost}_{jt} + \text{Brand}_j + \text{Store}_t + \text{YearMonth}_t + \xi^{(1)}_{jt} \tag{3.4}$$

$$\text{Price}_{jt} \times \text{cumPercOnlineSpend}_{it} = \gamma_2 \text{Cost}_{jt} + \gamma_3 \text{Cost}_{jt} \times \text{cumPercOnlineSpend}_{it} + \boldsymbol{\gamma}_4' \text{Control}_{ijt} + \xi^{(2)}_{ijt} \tag{3.5}$$

Here, Control_{ijt} is a vector of control variables (e.g., $\text{cumPercOnlineSpend}_{it}, \mathcal{I}\{\text{BuyPrevOcca}_{ijt}\}$, Customer_i, $\text{Brand}_j, \text{YearMonth}_t$). Next, following Petrin and Train (2010), the original error term in the utility specification, ϵ_{ijt}, can be decomposed as follows:

$$\epsilon_{ijt} = \lambda_1 \text{CF\_Price}_{jt} + \lambda_2 \text{CF\_Price\_cumPercOnlineSpend}_{ijt} + \tilde{\epsilon}_{ijt} \tag{3.6}$$

where $\tilde{\epsilon}_{ijt}$ is assumed to follow i.i.d. Type I extreme value distribution. We then include two

[12] Alternatively, some earlier work has suggested that a more parsimonious way to deal with the interaction term between an endogenous regressor and an exogenous variable is to estimate first-stage regression for price in Equation (3.4) and add the fitted residuals as an additional regressor in the model equation, without the need to add another first stage regression deal with the interaction term between price and online share (Ebbes et al., 2016). We also consider this alternative approach and find that the results remain consistent with our main findings. Please see Appendix §B.5.7 for details.

control functions into the utility equation as additional regressors, which is now given by:

$$\begin{aligned}U_{ijt} &= \alpha_0 \text{Price}_{jt} + \alpha_1 \text{cumPercOnlineSpend}_{it} + \alpha_2 \text{Price}_{jt} \times \text{cumPercOnlineSpend}_{it} \\ &+ \alpha_3 \mathcal{I}\{\text{BuyPrevOcca}_{ijt}\} + \delta_{ij} + \lambda_1 \text{CF\_Price}_{jt} + \lambda_2 \text{CF\_Price\_cumPercOnlineSpend}_{ijt} + \tilde{\epsilon}_{ijt} \\ &= \psi_{ijt} + \tilde{\epsilon}_{ijt} \quad \forall j \in \{1, \ldots, J\},\end{aligned} \tag{3.7}$$

where ψ_{ijt} is the deterministic part of the utility term and $\psi_{i0t} = 0$ for the outside option. Under the Type I extreme value assumption for $\tilde{\epsilon}_{ijt}$, the probability that consumer i chooses brand j at shopping occasion t is given by the following formula:

$$Pr_{ijt} = \frac{\exp(\psi_{ijt})}{\sum_{j'=0}^{J} \psi_{ij't}} \tag{3.8}$$

Let $\boldsymbol{\psi}_{it} = (\psi_{i0t}, \psi_{i1t}, \psi_{i2t}, \ldots, \psi_{iJt}) \in \mathbb{R}^{J+1}$ be the vector of nominal utilities and $\boldsymbol{y}_{it} = (y_{i0t}, y_{i1t}, \ldots, y_{iJt}) \in \mathbb{R}^{J+1}$ be the binary choice outcomes for consumer i at shopping occasion t, with $y_{ijt} = 1$ if consumer i chooses brand j at shopping occasion t and 0 otherwise. The likelihood of consumer i at purchase occasion t is given by:

$$l(\boldsymbol{\psi}_{it}; \boldsymbol{y}_{it}) = \prod_{j=0}^{J} (Pr_{ijt}(\boldsymbol{\psi}_{it}))^{y_{ijt}} \tag{3.9}$$

The log-likelihood function over all consumers and purchase occasions can be written as

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^{N} \sum_{t \in T_i} \log l(\boldsymbol{\psi}_{it}; \boldsymbol{y}_{it}) \tag{3.10}$$

where T_i is the set of purchase occasions associated with customer i. The parameter vector of interest is $\boldsymbol{\theta} = \{\alpha_0, \alpha_1, \alpha_2, \alpha_3, \lambda_1, \lambda_2, \{\delta_{ij}\}_{i=1,\ldots,N; j=1,\ldots,J}\}$. Note that this is a high-dimensional parameter vector – in addition to population parameters such as αs and λs, we also have customer-brand fixed-effects (δ_{ij}s), which scale with the sample size.

3.6.3 Estimation

Our main estimation strategy follows the standard two-step approach used in control function approaches; see (Petrin and Train, 2010) for details. However, the key additional challenge in our setting arises from the presence of high-dimensional customer-brand fixed effects. While high-dimensional fixed effects generally allow for more flexible substitution patterns (Jiang et al., 2025), doing so complicates the estimation procedure. We now discuss both the overview of our estimation strategy and also provide a summary of the solution concept that we adopt to address the high-dimensional fixed-effects.

First, we run OLS regressions for Price_{jt} in Equation (3.4) and $\text{Price}_{jt} \times \text{cumPercOnlineSpend}_{it}$ in Equation (3.5), respectively. We then obtain the residuals, CF_Price_{jt} and $\text{CF\_Price\_cumPercOnlineSpend}_{ijt}$, which serve as control functions. Second, we include these residuals as additional regressors in the logit model (i.e., Equation (3.7)). In a standard setting, estimation can proceed as usual at this stage. However, since we have a large number of customer-brand fixed effects (δ_{ij}), standard maximum likelihood estimation becomes computationally infeasible. To overcome this, we adopt a Minorization-Maximization (MM) algorithm for the logit model with fixed effects, an efficient estimation method proposed by Chen et al. (2022). The MM algorithm transforms the non-linear optimization problem (i.e., MLE estimation) into a sequence of simpler problems by iteratively constructing and maximizing a surrogate transfer minorization function $S(\cdot)$ that minorizes the log-likelihood (James, 2017). The transfer minorization function $S(\cdot)$ is less than the log-likelihood function everywhere except at the current best guess of the parameters, where it is equal. Chen et al. (2022) derives a transfer minorization function for the standard multinomial logit model in which there are no alternative-specific time-varying covariates in the utility function. In the theorem below, we extend their proof to cover a fully flexible logit model with covariates that may vary by individual, by alternative (brand), and over time.[13] We then derive the corresponding transfer minorization function $S(\cdot)$ for the log-likelihood $\mathcal{L}$.

Theorem. Let $\mathcal{L}(\boldsymbol{\theta})$ be the log-likelihood of the logit model, defined as in Equation (3.10).

[13]While our paper was the first to present this extension and proof for this specific case, in a recently posted update, Chen et al. (2025) present a similar extension.

Let $S(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$ be defined as:

$$S(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)}) = \mathcal{L}(\boldsymbol{\theta}^{(k)}) + \frac{1}{2}\sum_{i,j,t} h_j(\boldsymbol{\psi}_{it}^{(k)};\boldsymbol{y}_{it})^2 - \frac{1}{2}\sum_{i,j,t}[\psi_{ijt}^{(k)} + h_j(\boldsymbol{\psi}_{it}^{(k)};\boldsymbol{y}_{it}) - \psi_{ijt}]^2, \tag{3.11}$$

where

$$h_j(\boldsymbol{\psi}_{it}^{(k)};\boldsymbol{y}_{it}) = \frac{\partial \log l(\boldsymbol{\psi}_{it}^{(k)};\boldsymbol{y}_{it})}{\partial \psi_{ijt}^{(k)}} = y_{ijt} - Pr_{ijt}(\boldsymbol{\psi}_{it}^{(k)}) \tag{3.12}$$

$$\boldsymbol{\psi}_{it}^{(k)} = (\psi_{i0t}^{(k)}, \psi_{i1t}^{(k)}, \psi_{i2t}^{(k)}, \ldots, \psi_{iJt}^{(k)}) \tag{3.13}$$

Here, $h_j(\boldsymbol{\psi}_{it}^{(k)};\boldsymbol{y}_{it})$ is the j_{th} entry of the gradient vector of log-likelihood with respect to the nominal utility of alternative j, $\psi_{ijt}^{(k)}$. $\boldsymbol{\psi}_{it}^{(k)}$ denotes the vector of nominal utilities for consumer i at iteration k. Then, the function $S(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$ is a transfer minorization of $\mathcal{L}(\boldsymbol{\theta})$.

Proof. See Appendix B.3.1.

Based on the Theorem 1 and the proof in Chen et al. (2022), the iterative sequence $\boldsymbol{\theta}^{(k+1)} = \arg\max_\theta S(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$, $k = 0, 1, \ldots,$ converges to a local, if not global, maximum of $\mathcal{L}(\boldsymbol{\theta})$.[14]

Equation (3.11) is a second-order Taylor expansion of the log-likelihood function at $\boldsymbol{\theta}^{(k)}$. The model parameters we are searching over in the iterative MM algorithm ($\boldsymbol{\theta}$, rather than $\boldsymbol{\theta}^{(k)}$), appear only in the third term. Because the only term in $S(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$ that depends on $\boldsymbol{\theta}$ is the sum of squared residuals, maximizing $S(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$ is equivalent to minimizing this least squares objective, and the coefficients can be obtained by regressing $\psi_{ijt}^{(k)} + h_j(\boldsymbol{\psi}_{it}^{(k)};\boldsymbol{y}_{it})$ on the vector of explanatory variables $\boldsymbol{x}_{ijt} = \{\text{Price}_{jt}, \text{cumPercOnlineSpend}_{it}, \text{Price}_{jt} \times \text{cumPercOnlineSpend}_{it}, \mathcal{I}\{\text{BuyPrevOcca}_{ijt}\}, \text{CF\_Price}_{jt}, \text{CF\_Price\_cumPercOnlineSpend}_{ijt}, \delta_{ij}\}$ in the utility function. This linearization enables us to apply standard panel data methods despite the large number of fixed effects δ_{ij}. Moreover, another advantage of using the MM algorithm to estimate the logit model is that for customer-brand pairs with few or no purchases, maximum likelihood estimation formally drives δ_{ij} toward $-\infty$, so that the predicted choice probability is exactly zero. In practice, this boundary solution is numerically unstable.

[14]For more details, please refer to the proof of Theorem 1 in Chen et al. (2022), which shows the convergence of this iterative estimator to the true parameters.

The MM algorithm we employ circumvents this by iteratively solving a series of linearized regressions: never-chosen customer-brand pairs converge to large negative but finite estimates, which approximate the boundary solution while preserving numerical stability. Thus, the algorithm can include all (i, j) pairs without manual dropping, effectively regularizing the fixed effects associated with rate or absent choices.

In summary, the MM-based estimation algorithm that we employ is as follows:

- Initialization: Set the initial parameter estimates $\boldsymbol{\theta}^{(0)}$ including all customer-brand fixed effects as zero.
- Iterative Steps:
  1. Minorization step: Given $\boldsymbol{\theta}^{(k)}$, calculate $h_j(\boldsymbol{\psi}_{it}^{(k)}; \boldsymbol{y}_{it})$ and obtain $\psi_{ijt}^{(k)} + h_j(\boldsymbol{\psi}_{it}^{(k)}; \boldsymbol{y}_{it})$.
  2. Maximization step: Regress $\psi_{ijt}^{(k)} + h_j(\boldsymbol{\psi}_{it}^{(k)}; \boldsymbol{y}_{it})$ on the vector of variables $\boldsymbol{x}_{ijt}$ in the utility equation (3.7).
  3. Repeat until convergence, that is, $\|\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)}\|_\infty < 10^{-3}$.

Since the estimated residuals from the first stage enter the estimation in the second stage, the standard errors in the second step estimation tend to be downward biased, although the parameter estimates themselves are consistent (Wooldridge, 2010). Therefore, we use block bootstrapping with 500 replications, which treats individual customers as blocks. In each replication, we resample the same number of customers as in the original sample with replacement, re-estimate the demand model, and then obtain the bootstrapped standard errors for the demand parameters and price elasticities (Cameron and Trivedi, 2005).

3.6.4 Demand Model Parameter Estimates

Table 3.6 presents the estimation results of the demand model based on Equation (3.7). The standard errors of the parameter estimates are derived using block bootstrap with 500 replications. All the price coefficients are negative and statistically significant. The price interaction terms are negative and significant for three out of the four categories we study: dry dog food and for dog and cat hygiene categories, indicating that the adoption (and

Table 3.6. Coefficient Estimates of Demand Model - Main Results

Variables	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Price	-0.3159***	-0.1603***	-1.2464***	-0.221***
	(0.0235)	(0.0188)	(0.123)	(0.0209)
cumPercOnlineSpend	-0.3357	-1.4046	8.7512***	3.4655**
	(0.4601)	(1.0888)	(1.8675)	(1.0615)
Price * cumPercOnlineSpend	-0.1115**	0.018	-4.1601***	-0.284***
	(0.0421)	(0.091)	(0.8197)	(0.0713)
BuyPrevOcca	0.1512***	-0.5037***	-0.1821***	0.3247***
	(0.0273)	(0.0451)	(0.0284)	(0.0295)
CF_Price	0.0515	0.0468	0.1911	0.2805***
	(0.0293)	(0.0258)	(0.1815)	(0.0279)
CF_Price_cumPercOnlineSpend	0.0227	-0.0604	3.6335***	0.3006**
	(0.0803)	(0.1232)	(1.0007)	(0.0936)
Customer_Brand_FE	Yes	Yes	Yes	Yes
Observations	136,411	45,443	78,253	55,410
Log-likelihood	-84210.21	-30698.052	-55444.644	-37327.537

Bootstrapped standard-errors in parentheses via 500 replications

*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

extent of adoption) of a retailer's channels makes consumers *more* price sensitive during their subsequent offline shopping trips at that retailer's physical stores. The coefficient is insignificant for dry cat food. We discuss the potential sources for this heterogeneity in §3.7.1.

The magnitude of the coefficient of the interaction term in the logit model is difficult to interpret directly. Instead, we can interpret the magnitude of the effect we find by computing the marginal effects and price elasticities (Ai and Norton, 2003; Karaca-Mandic et al., 2012). To that end, we calculate price elasticities using the statistically significant coefficients, as specified in Equation (3.14) and discuss their implications in the next section.

3.7 ATT of Online Adoption on Offline Price Elasticity

Now that we have the demand parameter estimates and we see a negative and significant effect on the interaction term of interest, we measure the category-level own-price elasticities and formalize the ATT of the adoption of a retailer's online channels on offline price elasticities. In §3.7.1, we present the elasticity estimates from the estimation procedure. Next, in §3.7.2, we perform a within-customer comparison of adopters' offline price elasticity before and after the adoption of online shopping. In §3.7.3, we formalize the changes in offline price elasticity due to the adoption of online shopping using a staggered difference-in-differences (DID) model.

3.7.1 Elasticity Estimates

Given the demand model parameters, we can then derive a customer i's own-price elasticity for brand j in a purchase occasion t as shown below:

$$\eta_{ijt} = \frac{\partial Pr_{ijt}}{\partial \text{Price}_{jt}} \cdot \frac{\text{Price}_{jt}}{Pr_{ijt}} = (\alpha_0 + \alpha_2 \times \text{cumPercOnlineSpend}_{it}) \times \text{Price}_{jt} \times (1 - Pr_{ijt}) \quad (3.14)$$

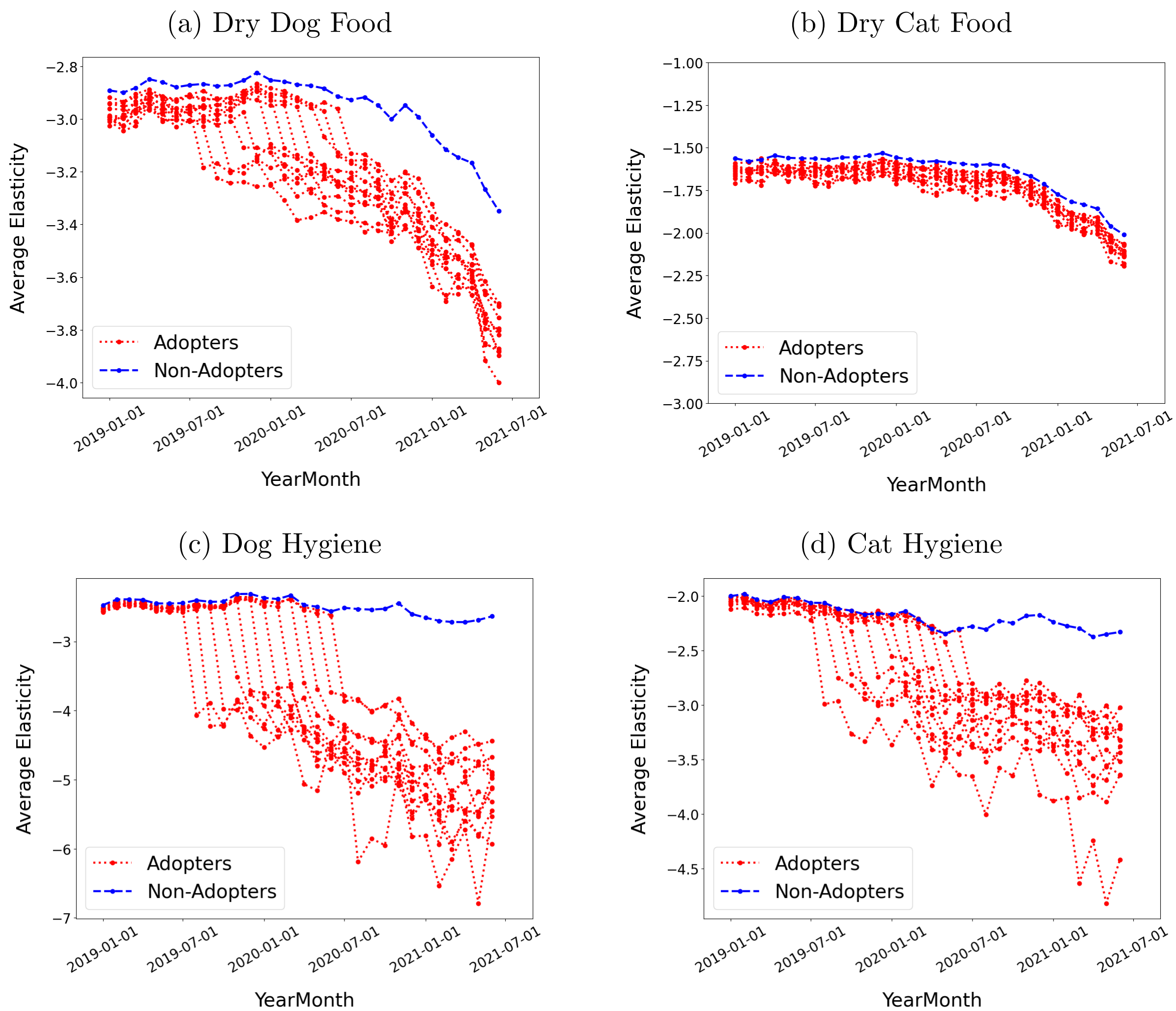
To estimate category-level own-price elasticity, we proceed in two steps. First, for each customer i, brand j in the customer-month purchase occasion t, we calculate the own-price elasticity η_{ijt} using Equation (3.14). We then average across the J brands to obtain a customer-month level own-price elasticity for each category, denoted by $\tilde{\eta}_{it} = \frac{1}{J}\sum_{j'} \eta_{ij't}$. Hereafter, we let t index the calendar time (i.e., year-month) rather than a purchase occasion. $\tilde{\eta}_{it}$ serves as the dependent variable in our subsequent regression analysis.

For easier interpretation, we further calculate the monthly average elasticity for each adopter and non-adopter cohort by taking the cross-sectional mean across all customers' own-price elasticities within that cohort. Figure 3.2 shows these monthly average offline price elasticities for both adopters and non-adopters of these product categories. The dashed line represents the baseline for non-adopters, while the twelve dotted lines represent adopter cohorts, each defined by the month in which customers first purchased online between July 2019 and June 2020.

Before online adoption, consumers' offline price elasticities closely mirror those of non-adopters, suggesting comparable pre-trends. However, after adopting online shopping, adopters consistently become more price sensitive in their offline shopping. The magnitude of the increase in offline price elasticity is most pronounced for dog hygiene and cat hygiene, moderate for dry dog food, and barely visible for dry cat food. The sharp increases in elasticities for the hygiene categories are likely due of the fact that these products have low switching costs: owners face little risk in trying a new hygienic mat for their dog if it is lower priced, with other attributes like pad size being standardized. The smaller changes to price elasticities in the two pet food categories, and a zero effect for cat food, are likely due to the fact that brand switching is not so simple in the case of pet food. Changing a dog's or cat's diet entails palatability trials, digestive risks, and rejection. Cats, in particular, are very picky eaters and have especially strong food preferences in terms of shape, smell, and texture (Nichelason, 2024). Since the main coefficient of interest in Table 3.6 for cat food is not statistically significant at the 5% level, and the elasticity figure also shows no effect for cat food, we drop this category in our subsequent analyses.

In sum, Figure 3.2 provides clear visual evidence supporting the idea that adoption of online shopping heightens consumers' offline price elasticity. Furthermore, this pattern holds true across all cohorts, as well as for the dry dog food and dog and cat hygiene categories. For

Figure 3.2. Average price elasticity between adopters (shown in twelve dotted lines) and non-adopters (shown in one dashed line) for each category.



cat food, we find a null effect; hence, we drop the category from our subsequent analyses. A drawback of the plots above is that we don't control for common time-specific shocks and/or unobserved customer heterogeneity. In the next few sections, we address these confounds using within-customer and staggered difference-in-differences (DID) analysis.

3.7.2 Within-Customer Analysis

We now quantify how adopters' offline price elasticity changes after they start shopping online at the focal retailer. We estimate two model specifications on the sample of adopters for each category:

$$\tilde{\eta}_{it} = \phi_0 + \phi_1 \times \mathcal{I}\{t \geq T_i^{\text{adopt}}\} + e_{it} \tag{3.15}$$

$$\tilde{\eta}_{it} = \text{Customer}_i + \text{YearMonth}_t + \phi_2 \times \mathcal{I}\{t \geq T_i^{\text{adopt}}\} + e_{it}, \tag{3.16}$$

where the binary indicator $\mathcal{I}\{t \geq T_i^{\text{adopt}}\}$ indicates the post-online-adoption period, which equals 1 if customer i has adopted online shopping at time t, and 0 otherwise. Customer_i and YearMonth_t are customer- and year-month fixed effects, which control for individual-specific time-invariant unobserved heterogeneity and time-specific unobserved shocks, respectively. Finally, because the outcome variable is an estimate (rather than data), we bootstrap standard errors. Specifically, we employ a block bootstrap with 500 replications, treating individual customers as blocks (as described in §3.6) to obtain standard errors.

Table 3.7 presents the parameter estimates (and bootstrapped standard errors) for these categories. We find that the coefficient on $\mathcal{I}\{t \geq T_i^{\text{adopt}}\}$ is consistently negative across all specifications and categories. This suggests that, on average, adopters become more sensitive in offline channels in these product categories after adopting online shopping (even after controlling for customer and time fixed effects).

Although this pre/post comparison isolates within-customer changes, it remains a before-after contrast confined to adopters. It does not control for common time-varying or macro shocks that could have affected the trend of price elasticities of both adopters and non-adopters. To effectively address this issue, we need to construct a counterfactual or control

Table 3.7. Within-Customer Analysis of Offline Price Elasticity - Main Results

Dependent Variable:	Offline Price Elasticity					
	Dry Dog Food		Dog Hygiene		Cat Hygiene	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	-2.935***		-2.464***		-2.125***	
	(0.2142)		(0.2429)		(0.1998)	
$\mathcal{I}\{t \geq T_i^{\text{adopt}}\}$	-0.3870***	-0.1173*	-1.916***	-0.7239**	-0.8860***	-0.3347**
	(0.0759)	(0.0543)	(0.3563)	(0.2361)	(0.1727)	(0.1250)
Fixed-effects						
Customer		Yes		Yes		Yes
YearMonth		Yes		Yes		Yes
Fit statistics						
Observations	29,416	29,416	28,375	28,375	15,081	15,081
R^2	0.31325	0.86657	0.34866	0.72137	0.37526	0.75980
Within R^2		0.04011		0.03352		0.03960

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

group of non-adopters, who were also offline customers of the firm in the pre-adoption period and continue to be offline only. To address this issue, we outline a difference-in-differences strategy in the next section.

3.7.3 Difference-In-Differences (DID) Analysis

To estimate the causal effect of interest – the average treatment effect on the treated (ATT), we use a difference-in-differences (DID) approach. This method compares adopters with non-adopters before and after online adoption. As a baseline specification, we estimate the following two-way fixed effects (TWFE) model:

$$\tilde{\eta}_{it} = \text{Customer}_i + \text{YearMonth}_t + \phi_{TWFE} \times \mathcal{I}\{\text{Adopter}_i\} \times \mathcal{I}\{t \geq T_i^{\text{adopt}}\} + e_{it}. \tag{3.17}$$

Here, ϕ_{TWFE} is our parameter of interest. $\mathcal{I}\{\text{Adopter}_i\}$ is a binary indicator that is 1 if customer i is an adopter and 0 otherwise. $\mathcal{I}\{t \geq T_i^{\text{adopt}}\}$ is a binary indicator equal to 1 if the year-month t is in the post-online adoption phase of customer i. We present the estimation results of the two-way fixed effects estimator in Table B.4 in Appendix §B.4. The estimates show that there is a significant increase in price elasticity for adopters post-adoption, as indicated by the negative and significant coefficient on ϕ_{TWFE} for these categories.

Although TWFE regression is widely used in empirical studies involving staggered adoption, it only provides consistent estimates of ATT under the assumption of homogeneous treatment effects, a condition that is rarely met in practice; see De Chaisemartin and d'Haultfoeuille (2020); Callaway and Sant'Anna (2021); Sun and Abraham (2021); Goodman-Bacon (2021) for detailed discussions of this issue. Specifically, Goodman-Bacon (2021) shows that the TWFE estimator (ϕ_{TWFE}) effectively represents a weighted average of all possible 2-by-2 combinations of pairwise difference-in-differences comparisons between groups of units treated at different times.

To address this issue, we adopt the staggered difference-in-differences (DID) framework proposed by Callaway and Sant'Anna (2021), hereafter referred to as the CS estimator as our main specification. The CS estimator addresses the biases associated with TWFE estimators by eliminating problematic comparisons between late-treated and already-treated units and has been applied recently in the marketing and economics literature (Ang, 2021; Braghieri et al., 2022; Bekkerman et al., 2023). This method is particularly suited to our context, because we have customers adopting online shopping over a one year period, in 12 cohorts (one cohort per month). As discussed earlier, because there were exogenous events during this time period (e.g., the onset of COVID-19), it is possible that the ATT is different across cohorts. Therefore, we use this estimator in our main analysis and then supplement the analysis with TWFE estimates as a robustness check (which provide similar results in terms of direction and magnitude). We refer the reader to the original paper for a discussion of the technical details of how the CS estimator works. The main idea is to identify the Average Treatment Effect on the Treated (ATT) for each post-treatment period, for consumers who are treated at different times, and then calculate the overall treatment effect as a weighted average of these ATTs. In our context, the ATT in month t for customers who adopt online shopping in month g can be defined as:

$$ATT(g,t) = \mathbb{E}[\tilde{\eta}_{it}(g) - \tilde{\eta}_{it}(0)|G_{ig} = 1], \tag{3.18}$$

where $\tilde{\eta}_{it}(g)$ is the observed elasticity of customer i at month t who adopts online shopping in month g. G_{ig} is a binary indicator which is set to one if the customer i adopts online shopping in month g. Since the adoption window in our study spans from July 2019 to June 2020, g ranges from 7 to 18. $\tilde{\eta}_{it}(0)$ represents the potential outcome (i.e. counterfactual elasticity) of customer i at month t, if they did not adopt online shopping. Under the "no anticipation of treatment" assumption from CS, $\tilde{\eta}_{it}(0) = \tilde{\eta}_{it}(g)$ for all $t < g$. In other words, the offline price elasticity of any customer i during the pre-treatment period does not depend on whether and when customer i adopts online shopping, and this represents the counterfactual elasticity if they had not adopted online shopping.

As discussed by Callaway and Sant'Anna (2021), this effect, $ATT(g,t)$, can be identified provided that parallel pre-trends hold. We provide a test of parallel pre-trends in Figure B.7 in Appendix §B.5.3.

The overall treatment effect (θ_o) is then calculated as a weighted average of $ATT(g,t)$ for all periods post-adoption across all groups:

$$\theta_o = \frac{1}{C_o} \sum_{g} \sum_{t>g} w_g ATT(g,t) \tag{3.19}$$

where w_g is the weight assigned to units in group g, which is the proportion of the number of online adopters in treatment group g relative to the total number of adopters. C_o normalizes the weights to sum to one.

Table 3.8. ATT of Online Adoption on Offline Price Elasticity Over All Post-Adoption Period (CS Estimator)

Dependent Variable:	Offline Price Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
ATT	-0.2482**	-1.7922***	-0.7381***
	(0.0945)	(0.3678)	(0.1794)
ATT Relative to Adopters' Pre-Adoption Elasticity	8.4%	72.7%	34.7%
Observations	136,411	78,253	55,410

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

We present the overall treatment effect estimates (θ_o) in Table 3.8. Since the customer-month level price elasticities are themselves estimates from our demand model with confidence intervals around them (and not directly observed from data), we compute standard errors for our ATT estimates using customer-level block bootstraps with 500 replications. Figure B.3 in Appendix §B.4 shows the bootstrapped distribution of these ATT estimates. The results in Table 3.8 suggest that consumers become more price sensitive in the offline channel after they adopt online shopping, and this finding is consistent across these categories. The effect is stronger for the dog and cat hygiene categories, consistent with the elasticity visualization in Figure 3.2. To help with interpretation, we calculate the relative magnitude of the elasticity increase in percentage terms by dividing the ATT by the average pre-adoption price elasticity of adopters and show these percentage changes in the second row of Table 3.8. The most substantial effects are observed for dog hygiene and cat hygiene categories; offline price elasticity increases by 72.7% (i.e., -1.792 / -2.464) for dog hygiene and 34.7% (i.e., -0.738 / -2.125) for cat hygiene. The increase in offline price elasticity is lower for dry dog food, at 8.4% (i.e., -0.248 / -2.935). As discussed in §3.7.1, the pronounced increase in offline price elasticity for hygiene products (vs. dry food products) is likely driven by the lower switching costs associated with trying new brands of hygiene mats or other hygiene-related items. In sum, consistent with the descriptive evidence earlier, we find that consumers become more price sensitive in their offline shopping trips when they start shopping online.

3.8 Generalizability and Robustness Checks

We now discuss the generalizability of our findings, provide some extensions, and present a series of tests and alternative specifications to establish the robustness of our findings.

3.8.1 Generalizability

We now discuss the generalizability of our findings and provide some evidence in favor of portability to other settings.

First, consumers may adopt online shopping for different reasons, and a natural question is whether our results are specific to consumers who adopt online shopping for certain reasons but not others. To examine this question, we leverage the fact that our observation period starts before the onset of the COVID-19 pandemic and continues through it. The cohorts that adopted from July 2019 to Feb 2020 are likely to have done so organically, while those who adopted from March 2020 are likely to have done so for exogenous reasons, i.e., the COVID-19 pandemic. Indeed, there is a big surge in online adoptions starting March 2020 (see Figure 3.1). We use this exogenous shock in the reason for adoption across cohorts to examine this issue. Overall, we find that the reason why a consumer switches to online shopping does not seem to have a significant impact on post-adoption price sensitivity: the ATT of cohorts that switched during the COVID-19 pandemic is similar to that of cohorts that switched organically before the onset of COVID-19 (see Figure 3.2). This suggests that the online adoption effect is not sensitive to the reasons that drive a consumer's decision to adopt online shopping.

Second, we examine the extent to which our results are likely to be driven by the fact that online prices are slightly lower than offline prices in our setting. To that end, we examine the magnitude of the price differences between online and offline channels in our study (see details in Appendix §B.5.1). We find that online prices are only 2–6% lower than offline in our data. Furthermore, the variation in prices explained by channel is significantly smaller than the variation in prices explained by SKU-level differences – across all categories, only 4% of the variation in prices is explained by channels, whereas when we include SKU fixed effects, this increases to 65%. Overall, these data patterns suggest that the differences between online and offline prices are relatively small and are unlikely to be the primary driver of our results. Other potential mechanisms, such as consumers becoming more accustomed to comparing prices across brands or their exposure to a wider variety of prices, are likely to be at play as well (as discussed in §3.1). Thus, we expect that our results are likely to generalize to other empirical settings where online prices are the same or higher than offline prices.[15]

Nevertheless, we note that we are unable to tease out the exact source of these effects because we lack granular data on consumers' search and shopping behavior in the online channel. We do not observe online click-stream data, so we cannot comment on price search behavior online or the variety of prices seen online. We also do not observe any search behavior in offline stores. Further, while there are some differences in offline prices across physical stores at any given point in time, we do not have a sufficiently large proportion of stores that are consistently higher or lower priced than others throughout our entire time period. Therefore, we cannot conduct a clean test of whether the effect we observe is driven

[15] In practice, there is considerable heterogeneity in whether online prices are lower, the same, or higher than offline prices across retailers, countries, and product categories (Cavallo, 2017).

more by price differences across the offline and online channels, the variety of prices that the consumer sees, or changes in search behavior that occur after adopting online shopping. We hope that future research can further dis-entangle which of these mechanisms are at play and to what extent. Further studies of such spillover effects in other countries, product categories, and retailers can also provide a more comprehensive understanding of the extent to which these effects manifest in different settings.

3.8.2 Robustness checks

We first examine alternative demand specifications in §3.8.2.1 and then consider matching methods and alternative control groups in §3.8.2.2 and 3.8.2.3, respectively. Finally, in §3.8.2.4, we consider an alternative definition of the purchase occasions.

3.8.2.1 Alternative Demand Specifications

We now consider two different demand model specifications:

- **Binary online adoption variable:** In our main demand model, we use the continuous variable
  $\text{cumPercOnlineSpend}_{it}$ to capture how consumers' offline price sensitivity changes after adopting online shopping. We now consider an alternative specification where we treat online adoption as a binary treatment. To that end, we use the binary indicator, $\text{Post\_OnlineAdoption}_{it}$, which takes value one if customer i has adopted online shopping by time t, and zero otherwise. We substitute $\text{cumPercOnlineSpend}_{it}$ with $\text{Post\_OnlineAdoption}_{it}$ in Equation (3.3), and also modify the corresponding control function term accordingly. We estimate revised demand model and compute the average elasticity at the customer-month level for each product category following the same steps as before. The details of the model and estimation results are provided in Appendix §B.5.2. The conclusions from this alternative demand model and the corresponding ATT estimates are consistent with our main results.

- **Accounting for a potential change in state dependence or brand loyalty due to online adoption:** In our main model, we include customer-brand fixed effects to capture time-invariant preferences and use $\mathcal{I}\{\text{BuyPrevOcca}_{ijt}\}$ to capture the state dependence. We now consider a specification with an additional control variable, $\mathcal{I}\{\text{BuyPrevOcca}_{ijt}\} \times \text{cumPercOnlineSpend}_{it}$, which captures the interaction between state dependence (e.g., repeated purchase of a brand) and the share of online spending. This can account for any effects that online adoption may have on brand loyalty.

The utility specifications and the estimation results from these two exercises are shown in Appendix §B.5.2. We find that the results are consistent with our main specification. We also do not find evidence of significant changes to offline brand loyalty after consumers adopt online shopping.

3.8.2.2 Matching and IPW Approaches to Control for Selection on Observables

Consumers self-select into adopting a retailer's online channel (or not), and as shown in Appendix §B.1.2, there are some systematic differences between adopters and non-adopters in terms of observable characteristics like demographics. In general, such differences are not an issue as long as the parallel pre-trends assumption holds (see Chapter 5 of Angrist and Pischke (2008) and Roth et al. (2023) for additional discussions). To that end, we consider the pre-adoption trends in price elasticities between adopters and non-adopters and verify that the parallel pre-trends assumption holds in Appendix §B.5.3.

Nevertheless, differences in characteristics between groups may still potentially lead to differences in price elasticities over time in some way that is not showing up in the first few months of data or is not being accounted for in our model. For instance, there may be some factor that affects younger customers who are more likely to adopt online shopping, and only shows up in the later part of our data period. To provide additional credibility to our findings and to ensure that they are robust to any time-varying factors correlated with observable characteristics, we use propensity score-based methods like Inverse Propensity Weighting (IPW) and Propensity Score Matching (PSM) to account for differences in observable characteristics between the groups, which improves the comparability between the groups. These methods improve covariate balance, discard poorly matched customers, and make our results more robust to model misspecification with respect to any time-varying factors that are correlated with observable characteristics.

We perform the propensity score estimation and matching for each adoption cohort separately. Specifically, we define the pre-adoption period separately for each cohort and measure pre-adoption variables (e.g., monthly average spend) using all purchase data up to the month of adoption (Gu and Kannan, 2021). For example, when identifying matched non-adopters for customers who adopt in July 2019, we measure the pre-adoption variables for both adopters and non-adopters using the data up to June 30, 2019, and estimate the propensity score model for the adoption cohort in July 2019. We repeat this process for each cohort and estimate twelve different propensity score models corresponding to each monthly adoption cohort. We compute propensity scores with a logistic regression model and the variables included are demographics (e.g., gender, age, and household income) and pre-adoption purchase variables (i.e. average monthly spend, number of unique orders, unique brands, and unique categories). With the estimated propensity scores, we use our CS estimator to estimate the ATT using both IPW and PSM methods. The propensity score model specification and detailed results from the IPW and PSM methods are shown in Appendix B.5.4. We find that our conclusions remain the same.

3.8.2.3 Alternative Control Groups

While we do our best to account for differences in observable characteristics between adopters and non-adopters in §3.8.2.2, it may be that there are certain time-specific unobserved factors that only affect adopters or would-be adopters, or disproportionately affect them more than non-adopters, thus potentially making non-adopters an unsuitable control group. We conduct an additional robustness check, leveraging the fact that we have multiple cohorts of consumers who adopted online shopping at different times. Following Narang and Shankar (2019) and Manchanda et al. (2015), we consider an alternative specification where later adopter cohorts serve as control groups for earlier adopter cohorts. Since all of these cohorts adopt online shopping at some point, the differences in any time-specific unobservables (that correlate with factors that drive the decision of whether a consumer chooses to adopt online shopping or not) should be minimized across these different adoption cohorts, at least relative to non-adopters. The results from this alternative CS estimator are shown in Appendix §B.5.5. Overall, we see that the conclusions remain similar to those from our main specification, i.e., that consumers become more offline price sensitive after adopting online shopping.

3.8.2.4 Alternative Definition of Purchase Occasions

In §3.3.2, we define a purchase occasion as any customer-month during which we observe any offline interaction with the retailer (similar to a week-store visit in Gordon et al. (2013)). In this section, we consider an alternative definition. Specifically, if a customer visits the retailer (i.e., either in-store or online) during a month and purchases in any category (not limited to the categories we focus on in this study), that customer-month is included as a purchase occasion for all four categories. If a customer makes no purchases in any category during the month, that customer-month is not included as a purchase occasion. In other words, a purchase occasion represents any customer-month during which we observe any interaction with the retailer, either offline or online.

The detailed demand model estimates and ATT estimation results are presented in Appendix §B.5.6. Overall, our conclusions remain similar to those from our main specification, i.e., that consumers become more offline price sensitive after adopting online shopping.

3.9 Managerial Implications and Conclusions

The rapid growth of e-commerce has fundamentally reshaped consumer behavior, but its spillover effects on offline shopping dynamics remain understudied. In this paper, we investigate how the adoption of online shopping using a retailer's e-commerce channels influences consumers' offline price sensitivity when shopping at that retailer's physical stores, utilizing transaction-level data from a multichannel pet supplies retailer in Brazil. We find that consumers who adopt online shopping become more price sensitive in their offline purchases. The relative increase in offline price elasticity, defined as the ATT relative to the pre-adoption average price elasticity of the treated group, is quite heterogeneous across categories, with a 72.7% and a 34.7% increase for dog hygiene and cat hygiene categories, and a smaller increase of 8.4% for dog food. We find no effect on cat food, presumably because it is hard to switch brands easily in this category.

From a substantive standpoint, our results contribute to the literature on price elasticity by identifying online adoption as a novel determinant of offline price sensitivity. This factor has so far been ignored in the demand estimation literature, which usually only accounts for factors like consumer demographics and macro-economic shocks to affect consumer price sensitivity. From a managerial standpoint, our findings have implications for retailer offline pricing and profitability, emphasizing the need for multichannel retailers to account for the online adoption effect (i.e., heightened offline price sensitivity when consumers adopt their online channels) in their pricing decisions. Further, our approach can be used to personalize prices depending on customers' channel adoption status. While existing literature has largely focused on pricing and personalization based on customer demographics and prior purchase behavior (Rossi and Allenby, 1993; Chen et al., 2001; Jiang et al., 2021; Jain et al., 2024), our work suggests that incorporating cross-channel effects into measures of price elasticity and pricing strategy can further enhance profits.

Finally, our paper opens many avenues for future research. First, future research could explore whether these effects generalize to other markets (e.g., non-discretionary goods or regions with lower e-commerce penetration) and investigate mediating factors, such as the role of mobile price-checking apps or subscription models. Further, while our paper documents one aspect of consumer behavior (price sensitivity), multi-channel shopping can induce other spillover effects from one channel to another. For instance, product categories purchased offline may change after adopting online shopping (Huyghe et al., 2017), and there may be cross-media effects between channels (Joo et al., 2014; Du et al., 2019). Future research can investigate other such spillover effects and their effect on cross-channel purchases. Ultimately, as retail continues to evolve toward omnichannel integration, understanding these interdependencies will be critical for both academics and firms aiming to thrive in a digitally transformed landscape.

BIBLIOGRAPHY

Ai, Chunrong and Edward C Norton (2003), 'Interaction terms in logit and probit models', *Economics letters* **80**(1), 123–129.

Allenby, Greg M and Peter E Rossi (1998), 'Marketing models of consumer heterogeneity', *Journal of econometrics* **89**(1-2), 57–78.

Andrews, Rick L and Imran S Currim (2004), 'Behavioural differences between consumers attracted to shopping online versus traditional supermarkets: implications for enterprise design and marketing strategy', *International Journal of Internet Marketing and Advertising* **1**(1), 38–61.

Ang, Desmond (2021), 'The effects of police violence on inner-city students', *The Quarterly Journal of Economics* **136**(1), 115–168.

Angrist, Joshua D and Guido W Imbens (1995), 'Two-stage least squares estimation of average causal effects in models with variable treatment intensity', *Journal of the American statistical Association* **90**(430), 431–442.

Angrist, Joshua and Joern-Steffen Pischke (2008), 'Mostly harmless econometrics (princeton uni'.

Ansari, Asim, Carl F Mela and Scott A Neslin (2008), 'Customer channel migration', *Journal of Marketing Research* **45**(1), 60–76.

Aparicio, Diego, Dean Eckles and Madhav Kumar (2024), 'Algorithmic pricing and consumer sensitivity to price variability'.

Autor, David H (2003), 'Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing', *Journal of labor economics* **21**(1), 1–42.

Avery, Jill, Thomas J Steenburgh, John Deighton and Mary Caravella (2012), 'Adding bricks to clicks: Predicting the patterns of cross-channel elasticities over time', *Journal of Marketing* **76**(3), 96–111.

Baker, Andrew C, David F Larcker and Charles CY Wang (2022), 'How much should we trust staggered difference-in-differences estimates?', *Journal of Financial Economics* **144**(2), 370–395.

Bawa, Kapil and Robert W Shoemaker (1987), 'The coupon-prone consumer: some findings based on purchase behavior across product classes', *Journal of marketing* **51**(4), 99–110.

Bawa, Kapil and Robert W Shoemaker (1989), 'Analyzing incremental sales from a direct mail coupon promotion', *Journal of marketing* **53**(3), 66–78.

Bekkerman, Ron, Maxime C Cohen, Edward Kung, John Maiden and Davide Proserpio (2023), 'The effect of short-term rentals on residential investment', *Marketing Science* **42**(4), 819–834.

Bell, David R, Santiago Gallino and Antonio Moreno (2018), 'Offline showrooms in omnichannel retail: Demand and operational benefits', *Management Science* **64**(4), 1629–1651.

Berry, Steven T (1994), 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics* pp. 242–262.

Berry, Steven T, James A Levinsohn and Ariel Pakes (1995), 'Automobile prices in market equilibrium'.

Berry, Steven T and Philip A Haile (2021), Foundations of demand estimation, *in* 'Handbook of industrial organization', Vol. 4, Elsevier, pp. 1–62.

Bijmolt, Tammo HA, Harald J Van Heerde and Rik GM Pieters (2005), 'New empirical generalizations on the determinants of price elasticity', *Journal of marketing research* **42**(2), 141–156.

Biswas, Shirsho, Hema Yoganarasimhan and Haonan Zhang (2024), 'Channel choice and customer value', *Available at SSRN 4747756* .

Biswas, Shirsho, Hema Yoganarasimhan and Haonan Zhang (2025*a*), 'Channel adoption pathways and post-adoption behavior', *Available at SSRN 4747756* .

Biswas, Shirsho, Hema Yoganarasimhan and Haonan Zhang (2025*b*), 'How does online shopping affect offline price sensitivity?', *arXiv preprint arXiv:2506.15103* .

Bloomberg (2020), 'Digital transformation outlook', *Bloomberg Intelligence* .

Bradlow, Eric (1998), 'Encouragement designs: an approach to self-selected samples in an experimental design', *Marketing Letters* **9**(4), 383–391.

Braghieri, Luca, Ro'ee Levy and Alexey Makarin (2022), 'Social media and mental health', *American Economic Review* **112**(11), 3660–3693.

Brazil Institute of Geography and Statistics (IBGE) (2024), 'Brazil's extended national consumer index (ipca)'. Accessed: October 1, 2024.
URL: *https://www.ibge.gov.br/en/statistics/economic/prices-and-costs/*

Brynjolfsson, Erik and Michael D Smith (2000), 'Frictionless commerce? a comparison of internet and conventional retailers', *Management science* **46**(4), 563–585.

Callaway, Brantly and Pedro HC Sant'Anna (2021), 'Difference-in-differences with multiple time periods', *Journal of Econometrics* **225**(2), 200–230.

Cameron, A Colin and Pravin K Trivedi (2005), *Microeconometrics: methods and applications*, Cambridge university press.

Cavallo, Alberto (2017), 'Are online and offline prices similar? evidence from large multi-channel retailers', *American Economic Review* **107**(1), 283–303.

Chen, Guangying, Cheng Lu, Tat Chan, Zhengling Qi, Dennis Zhang, Yaran Jin and Miaoyu Yang (2024), 'A new estimator for randomized controlled trials with non-compliance: Theory and empirical evidence', *Available at SSRN 4864490* .

Chen, Mingli, Ian Meeker, Marc Rysman and Shuang Wang (2025), 'An mm algorithm for fixed effects multinomial logit models', *Unpublished manuscript, Boston University* .

Chen, Mingli, Marc Rysman, Shuang Wang and Krzysztof Wozniak (2022), An mm algorithm for fixed effects in multinomial models with an application to payment choice in grocery stores, Technical report, Working Papers.

Chen, Yuxin, Chakravarthi Narasimhan and Z John Zhang (2001), 'Individual marketing with imperfect targetability', *Marketing Science* **20**(1), 23–41.

Chib, Siddhartha, PB Seetharaman and Andrei Strijnev (2004), 'Model of brand choice with a no-purchase option calibrated to scanner-panel data', *Journal of marketing research* **41**(2), 184–196.

Chintagunta, Pradeep, Jean-Pierre Dubé and Khim Yong Goh (2005), 'Beyond the endogeneity bias: The effect of unmeasured brand characteristics on household-level brand choice models', *Management Science* **51**(5), 832–849.

Chintagunta, Pradeep K, Dipak C Jain and Naufel J Vilcassim (1991), 'Investigating heterogeneity in brand preferences in logit models for panel data', *Journal of Marketing Research* **28**(4), 417–428.

Chintala, Sai Chand, Jūra Liaukonytė and Nathan Yang (2023), 'Browsing the aisles or browsing the app? how online grocery shopping is changing what we buy', *Marketing Science* .

Chu, Junhong, Marta Arce-Urriza, José-Javier Cebollada-Calvo and Pradeep K Chintagunta (2010), 'An empirical analysis of shopping behavior across online and offline channels for grocery products: The moderating effects of household and product characteristics', *Journal of Interactive Marketing* **24**(4), 251–268.

Chu, Junhong, Pradeep Chintagunta and Javier Cebollada (2008), 'Research note—a comparison of within-household price sensitivity across online and offline channels', *Marketing science* **27**(2), 283–299.

Cooper, James C (2006), 'Prices and price dispersion in online and offline markets for contact lenses', *FTC Bureau of Economics working paper* (283).

Cui, Tony Haitao, Anindya Ghose, Hanna Halaburda, Raghuram Iyengar, Koen Pauwels, S Sriram, Catherine Tucker and Sriraman Venkataraman (2021), 'Informational challenges in omnichannel marketing: Remedies and future research', *Journal of Marketing* **85**(1), 103–120.

Danaher, Peter J, Isaac W Wilson and Robert A Davis (2003), 'A comparison of online and offline consumer brand loyalty', *Marketing Science* **22**(4), 461–476.

De Chaisemartin, Clément and Xavier d'Haultfoeuille (2020), 'Two-way fixed effects estimators with heterogeneous treatment effects', *American economic review* **110**(9), 2964–2996.

Degeratu, Alexandru M, Arvind Rangaswamy and Jianan Wu (2000), 'Consumer choice behavior in online and traditional supermarkets: The effects of brand name, price, and other search attributes', *International Journal of research in Marketing* **17**(1), 55–78.

Deloitte (2012), 'Omnichannel shoppers spend 1.5x more than single channel buyers', *Retail Touchpoints* .

Du, Rex Yuxing, Linli Xu and Kenneth C Wilbur (2019), 'Immediate responses of online brand search and price search to tv ads', *Journal of Marketing* **83**(4), 81–100.

Dubé, Jean-Pierre, Günter J Hitsch and Peter E Rossi (2010), 'State dependence and alternative explanations for consumer inertia', *The RAND Journal of Economics* **41**(3), 417–445.

Ebbes, Peter, Dominik Papies and Harald J van Heerde (2016), Dealing with endogeneity: A nontechnical guide for marketing researchers, *in* 'Handbook of market research', Springer, pp. 1–37.

Erdem, Tülin and Michael P Keane (1996), 'Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets', *Marketing science* **15**(1), 1–20.

Fader, Peter S, Bruce GS Hardie and Jen Shang (2010), 'Customer-base analysis in a discrete-time noncontractual setting', *Marketing Science* **29**(6), 1086–1108.

Fader, Peter S, Bruce GS Hardie and Ka Lok Lee (2005), 'Rfm and clv: Using iso-value curves for customer base analysis', *Journal of marketing research* **42**(4), 415–430.

Faria de Moura Villela, Edlaine, Rossana Verónica Mendoza López, Ana Paula Sayuri Sato, Fábio Morato de Oliveira, Eliseu Alves Waldman, Rafael Van den Bergh, Joseph Nelson Siewe Fodjo and Robert Colebunders (2021), 'Covid-19 outbreak in brazil: adherence to national preventive measures and impact on people's lives, an online survey', *BMC Public Health* **21**(1), 1–10.

Forman, Chris, Anindya Ghose and Avi Goldfarb (2009), 'Competition between local and electronic markets: How the benefit of buying online depends on where you live', *Management science* **55**(1), 47–57.

Fortune (2024), 'Pet care market', *Fortune Business Insights* .

Goodman-Bacon, Andrew (2021), 'Difference-in-differences with variation in treatment timing', *Journal of econometrics* **225**(2), 254–277.

Gordon, Brett R, Avi Goldfarb and Yang Li (2013), 'Does price elasticity vary with economic growth? a cross-category analysis', *Journal of Marketing Research* **50**(1), 4–23.

Goyal, Poonam (2023), 'E-commerce surge can add 1 trillion to us retail sales by 2027', *Bloomberg Intelligence* .

Gu, Xian and PK Kannan (2021), 'The dark side of mobile app adoption: Examining the impact on customers' multichannel purchase', *Journal of Marketing Research* **58**(2), 246–264.

Guadagni, Peter M and John DC Little (1983), 'A logit model of brand choice calibrated on scanner data', *Marketing science* **2**(3), 203–238.

Guadagni, Peter M and John DC Little (2008), 'A logit model of brand choice calibrated on scanner data', *Marketing Science* **27**(1), 29–48.

Gupta, Sunil, Dominique Hanssens, Bruce Hardie, Wiliam Kahn, V Kumar, Nathaniel Lin, Nalini Ravishanker and Srinivasaraghavan Sriram (2006), 'Modeling customer lifetime value', *Journal of service research* **9**(2), 139–155.

Haleem, Abbas (2025), 'Us ecommerce sales in 2024 more than double those of 2019'. Accessed: May 19, 2025.
URL: *https://www.digitalcommerce360.com/article/us-ecommerce-sales/*

Hendel, Igal and Aviv Nevo (2006), 'Sales and consumer inventory', *The RAND Journal of Economics* **37**(3), 543–561.

Hitsch, Günter J, Ali Hortacsu and Xiliang Lin (2021), 'Prices and promotions in us retail markets', *Quantitative Marketing and Economics* **19**(3), 289–368.

Hoch, Stephen J, Byung-Do Kim, Alan L Montgomery and Peter E Rossi (1995), 'Determinants of store-level price elasticity', *Journal of marketing Research* **32**(1), 17–29.

Huyghe, Elke, Julie Verstraeten, Maggie Geuens and Anneleen Van Kerckhove (2017), 'Clicks as a healthy alternative to bricks: how online grocery shopping reduces vice purchases', *Journal of Marketing Research* **54**(1), 61–74.

Jain, Lalit, Zhaoqi Li, Erfan Loghmani, Blake Mason and Hema Yoganarasimhan (2024), 'Effective adaptive exploration of prices and promotions in choice-based demand models', *Marketing Science* .

James, Jonathan (2017), 'Mm algorithm for general mixed multinomial logit models', *Journal of Applied Econometrics* **32**(4), 841–857.

Jedidi, Kamel, Carl F Mela and Sunil Gupta (1999), 'Managing advertising and promotion for long-run profitability', *Marketing science* **18**(1), 1–22.

Jiang, Zhaohui, Jun Li and Dennis Zhang (2025), 'A high-dimensional choice model for online retailing', *Management Science* **71**(4), 3320–3339.

Jiang, Zhenling, Tat Chan, Hai Che and Youwei Wang (2021), 'Consumer search and purchase: An empirical investigation of retargeting based on consumer online behaviors', *Marketing Science* **40**(2), 219–240.

Joo, Mingyu, Dinesh K Gauri and Kenneth C Wilbur (2020), 'Temporal distance and price responsiveness: Empirical investigation of the cruise industry', *Management Science* **66**(11), 5362–5388.

Joo, Mingyu, Kenneth C Wilbur, Bo Cowgill and Yi Zhu (2014), 'Television advertising and online search', *Management Science* **60**(1), 56–73.

Kalra, Ajay and Ronald C Goodstein (1998), 'The impact of advertising positioning strategies on consumer price sensitivity', *Journal of marketing research* **35**(2), 210–224.

Kamakura, Wagner A and Gary J Russell (1989), 'A probabilistic choice model for market segmentation and elasticity structure', *Journal of marketing research* **26**(4), 379–390.

Kanetkar, Vinay, Charles B Weinberg and Doyle L Weiss (1992), 'Price sensitivity and television advertising exposures: Some empirical findings', *Marketing Science* **11**(4), 359–371.

Karaca-Mandic, Pinar, Edward C Norton and Bryan Dowd (2012), 'Interaction terms in nonlinear models', *Health services research* **47**(1pt1), 255–274.

Kaul, Anil and Dick R Wittink (1995), 'Empirical generalizations about the impact of advertising on price sensitivity and price', *Marketing Science* **14**(3_supplement), G151–G160.

Kim, Jaehwan, Greg M Allenby and Peter E Rossi (2002), 'Modeling consumer demand for variety', *Marketing Science* **21**(3), 229–250.

Kopalle, Praveen K, Carl F Mela and Lawrence Marsh (1999), 'The dynamic effect of discounting on sales: Empirical analysis and normative pricing implications', *Marketing Science* **18**(3), 317–332.

Krishnamurthi, Lakshman and SP Raj (1988), 'A model of brand choice and purchase quantity price sensitivities', *Marketing Science* **7**(1), 1–20.

Kumar, Vipin and Rajkumar Venkatesan (2005), 'Who are the multichannel shoppers and how do they perform?: Correlates of multichannel shopping behavior', *Journal of Interactive Marketing* **19**(2), 44–62.

Kushwaha, Tarun and Venkatesh Shankar (2013), 'Are multichannel customers really more valuable? the moderating role of product category characteristics', *Journal of Marketing* **77**(4), 67–85.

Li, Hui (2021), 'Are e-books a different channel? multichannel management of digital products', *Quantitative Marketing and Economics* **19**(2), 179–225.

Manchanda, Puneet, Grant Packard and Adithya Pattabhiramaiah (2015), 'Social dollars: The economic impact of customer participation in a firm-sponsored online customer community', *Marketing Science* **34**(3), 367–387.

McKinsey & Company (2023), 'The next horizon for grocery e-commerce'. Accessed: 2024-09-29.
URL: *https://www.mckinsey.com/industries/retail/our-insights/the-next-horizon-for-grocery-e-commerce*

Mela, Carl F, Sunil Gupta and Donald R Lehmann (1997), 'The long-term impact of promotion and advertising on consumer brand choice', *Journal of Marketing research* **34**(2), 248–261.

Montaguti, Elisa, Scott A Neslin and Sara Valentini (2016), 'Can marketing campaigns induce multichannel buying and more profitable customers? a field experiment', *Marketing Science* **35**(2), 201–217.

Mummalaneni, Simha, Hema Yoganarasimhan and Varad Pathak (2023), 'How do content producers respond to engagement on social media platforms?'.

Mummalaneni, Simha, Hema Yoganarasimhan and Varad Pathak (2025), 'How do content producers respond to engagement on social media platforms?'.

Narang, Unnati and Venkatesh Shankar (2019), 'Mobile app introduction and online and offline purchases and product returns', *Marketing Science* **38**(5), 756–772.

Narasimhan, Chakravarthi (1984), 'A price discrimination theory of coupons', *Marketing Science* **3**(2), 128–147.

Neslin, Scott A (2022), 'The omnichannel continuum: Integrating online and offline channels along the customer journey', *Journal of retailing* **98**(1), 111–132.

Neslin, Scott A, Caroline Henderson and John Quelch (1985), 'Consumer promotions and the acceleration of product purchases', *Marketing Science* **4**(2), 147–165.

Neslin, Scott A, Dhruv Grewal, Robert Leghorn, Venkatesh Shankar, Marije L Teerling, Jacquelyn S Thomas and Peter C Verhoef (2006), 'Challenges and opportunities in multichannel customer management', *Journal of Service Research* **9**(2), 95–112.

Neslin, Scott A and Venkatesh Shankar (2009), 'Key issues in multichannel customer management: current knowledge and future directions', *Journal of Interactive Marketing* **23**(1), 70–81.

Nevo, Aviv (2001), 'Measuring market power in the ready-to-eat cereal industry', *Econometrica* **69**(2), 307–342.

Nichelason, Amy (2024), 'Expert q&a: Understanding pet food'. Accessed: May 21, 2025.
URL: *https://www.vetmed.wisc.edu/expert-qa-understanding-pet-food/*

Petrin, Amil and Kenneth Train (2010), 'A control function approach to endogeneity in consumer choice models', *Journal of marketing research* **47**(1), 3–13.

Pozzi, Andrea (2012), 'Shopping cost and brand exploration in online grocery', *American Economic Journal: Microeconomics* **4**(3), 96–120.

PRNewswire (2023), 'Omnichannel shoppers spend 1.5x more than single channel buyers', *PRNewswire* .

Rambachan, Ashesh and Jonathan Roth (2023), 'A more credible approach to parallel trends', *Review of Economic Studies* **90**(5), 2555–2591.

Rosenbaum, Paul R and Donald B Rubin (1985), 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score', *The American Statistician* **39**(1), 33–38.

Rossi, Peter E and Greg M Allenby (1993), 'A bayesian approach to estimating household parameters', *Journal of Marketing Research* **30**(2), 171–182.

Rossi, Peter E, Robert E McCulloch and Greg M Allenby (1996), 'The value of purchase history data in target marketing', *Marketing Science* **15**(4), 321–340.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski and John Poe (2023), 'What's trending in difference-in-differences? a synthesis of the recent econometrics literature', *Journal of Econometrics* **235**(2), 2218–2244.

Sant'Anna, Pedro HC and Jun Zhao (2020), 'Doubly robust difference-in-differences estimators', *Journal of Econometrics* **219**(1), 101–122.

Shriver, Scott K and Bryan Bollinger (2022), 'Demand expansion and cannibalization effects from retail store entry: A structural analysis of multichannel demand', *Management Science* **68**(12), 8829–8856.

Stuart, Elizabeth A, Gary King, Kosuke Imai and Daniel Ho (2011), 'Matchit: nonparametric preprocessing for parametric causal inference', *Journal of statistical software* .

Sun, Liyang and Sarah Abraham (2021), 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of econometrics* **225**(2), 175–199.

Syrgkanis, Vasilis, Victor Lei, Miruna Oprescu, Maggie Hei, Keith Battocchi and Greg Lewis (2019), 'Machine learning estimation of heterogeneous treatment effects with instruments', *Advances in Neural Information Processing Systems* **32**.

Tellis, Gerard J (1988), 'The price elasticity of selective demand: A meta-analysis of econometric models of sales', *Journal of marketing research* **25**(4), 331–341.

Thomas, Jacquelyn S and Ursula Y Sullivan (2005), 'Managing marketing communications with multichannel customers', *Journal of Marketing* **69**(4), 239–251.

Van Heerde, Harald J and Scott A Neslin (2017), 'Sales promotion models', *Handbook of Marketing Decision Models* pp. 13–77.

Venkatesan, Rajkumar and Nalini Ravishanker (2007), 'Multichannel shopping: causes and consequences', *Journal of Marketing* **71**(2), 114–132.

Verhoef, Peter C (2012), 'Multichannel customer management strategy', *Handbook of Marketing Strategy* pp. 135–150.

Verplanken, Bas and Wendy Wood (2006), 'Interventions to break and create consumer habits', *Journal of public policy & marketing* **25**(1), 90–103.

Villas-Boas, J Miguel and Russell S Winer (1999), 'Endogeneity in brand choice models', *Management science* **45**(10), 1324–1338.

Wang, Kitty and Avi Goldfarb (2017), 'Can offline stores drive online sales?', *Journal of Marketing Research* **54**(5), 706–719.

Wang, Rebecca Jen-Hui, Edward C Malthouse and Lakshman Krishnamurthi (2015), 'On the go: How mobile shopping affects customer purchase behavior', *Journal of Retailing* **91**(2), 217–234.

Wood, Wendy and David T Neal (2009), 'The habitual consumer', *Journal of Consumer Psychology* **19**(4), 579–592.

Wooldridge, Jeffrey M (2010), *Econometric analysis of cross section and panel data*, MIT press.

Wooldridge, Jeffrey M (2015), 'Control function methods in applied econometrics', *Journal of Human Resources* **50**(2), 420–445.

Yang, Botao and Andrew T Ching (2014), 'Dynamics of consumer adoption of financial innovation: The case of atm cards', *Management Science* **60**(4), 903–922.

Yu, Qi, Ron Berman and Eric Bradlow (2021), 'The dark side of category expansion: Will (and which) existing ones' pay the price'?', *Available at SSRN 3837014* .

Zhang, Dennis J, Hengchen Dai, Lingxiu Dong, Qian Wu, Lifan Guo and Xiaofei Liu (2019), 'The value of pop-up stores on retailing platforms: Evidence from a field experiment with alibaba', *Management Science* **65**(11), 5142–5151.

Appendix A

APPENDIX TO CHAPTER 2

A.1 Data cleaning

Table A.1. Item categories with substantially different availability time via online and offline channels (YY-MM-DD)

Item Category	First Offline Purchase Date	First Online Purchase Date
Large size food	2020-09-12	2020-02-28
Human food	2019-01-02	2021-02-12
Large size pharmacy	2021-02-03	2020-12-09
Pet Cleaning Supplies	2019-01-03	2022-03-13

There are four product categories with significant discrepancies in their initial availability time on the offline and online channels. For example, the product category *Large size food* (e.g. pack size greater than 10KG) was available online in February 2020 but only became available offline in September 2020. In addition, the product category Pet Cleaning Supplies was only available offline before COVID-19 but became available online in August 2022.

Including these food categories in our analysis can be problematic since it also reflects retailers' assortment and the channel-specific effects (i.e. consumers can only buy some products exclusively from one channel). Therefore, we exclude these four categories from our analysis. The fraction of sales of these four categories only accounts for 0.63% of the total revenue during the sample period, and as such unlikely to have any meaningful impact on our findings.

A.2 Appendix to Summary Statistics

In this section, we summarize the demographics, pre-adoption, and post-adoption purchase behaviors of three customer groups for each focal event: *event adopters*, *organic adopters*, and *offline-only customers*.

To characterize the pre-online-adoption purchase behavior of each group of customers, we compute several consumer-level metrics, such as *AvgSpendPerMonth*, *AvgOrdersPerMonth*, *AvgUniqueCategoriesPerMonth*, using data from the pre-adoption period specific to each case study. We summarize these variables for each group – *event adopters*, *organic adopters*, and *offline-only* – by calculating the group mean as follows: $\frac{1}{N_g}\sum_{i=1}^{N_g}(\frac{1}{T_i^{pre}}\sum_{t=1}^{T_i^{pre}} X_{it})$, where i denotes a customer belonging to group g, t denotes the year-month index, T_i^{pre} denotes the length of the consumer i's panel during the pre-adoption period, and N_g is the number of customers in group g. We next summarize the post-online adoption behaviors in a similar way using the post-online adoption data under each case study. The post-online adoption means for each group g – *event adopters*, *organic adopters*, and *offline-only* – are calculated as follows: $\frac{1}{N_g}\sum_{i=1}^{N_g}(\frac{1}{T_i^{post}}\sum_{t=1}^{T_i^{post}} X_{it})$, where i denotes a customer belonging to group g, t denotes the year-month index, T_i^{post} denotes the length of the consumer i's panel during the post-adoption period, and N_g is the number of customers in group g.

A.2.1 Event 1: COVID-19

In this section, we present the descriptive statistics for customers in the sample for the COVID-19, supplementing those in §2.5.3. We first summarize demographics in Table A.2. We also report customer-level summary statistics for average purchase variables during the post-adoption period in Tables A.3, complementing the pre-adoption statistics shown in Table 2.4 in §2.5.3. Finally, in Table A.4, we present more detailed summary statistics (e.g., median and quartiles) for all three groups during the post-online-adoption periods (i.e., between May 1st, 2020 and April 30, 2023). Note that the counts of *organic adopters* and *covid adopters* from Table A.4 slightly differ from the raw count based on our cohorts' definition because 469 (3.66% of) *organic adopters* and 933 (2.81% of) *covid adopters* churned before 2020-05-01, and therefore are not observed in the post-COVID-19 panel. Including these customers does not affect our identification of post-online-adoption ATT, and the estimation results remain the same.

Table A.2. Customer Demographics – Summary Statistics of Gender (Panel A) and Geographic State (Panel B).

Categories	*Covid Adopter* (%)	*Organic Adopter* (%)	*Offline-only* (%)
Panel A: Gender			
No Information	12.09	13.52	15.63
Female	59.59	57.36	46.51
Male	28.31	29.13	37.86
Total	100.00	100.00	100.00
Panel B: Geographic State			
São Paulo	58.78	49.68	63.70
Minas Gerais	8.58	12.13	4.81
Rio de Janeiro	6.69	6.95	4.30
Distrito Federal	6.21	5.48	4.26
Zip_State_Missing	4.07	4.28	10.34
Santa Catarina	3.33	3.46	2.90
Goiás	2.37	3.36	2.13
Bahia	1.74	3.31	0.69
Rio Grande do Sul	3.21	2.77	2.20
Espírito Santo	1.27	2.43	0.91
Mato Grosso do Sul	0.93	1.88	1.05
Paraná	1.68	1.86	2.31
Other States	1.14	2.41	0.39
Total	100.00	100.00	100.00

A.2.2 Event 2: Black Friday

In this section, we present the descriptive statistics for customers in the sample for the Black Friday event, supplementing those in §2.5.3. We first summarize demographics in Table A.5. We also report customer-level summary statistics for average purchase variables during

Table A.3. Customer-Level Summary Statistics of **Post-online-adoption** Purchase Behaviors between May 1st 2020 and April 30th 2023 (Panel A). Columns 1 - 2 describe statistics for *Covid adopters* and *Organic adopters*, respectively. Column 3 calculates the difference between *Covid adopters* and *Organic adopters*, and columns 4 and 5 show the two-sample t-test and p-value results.

Variable	*Covid Adopters* (Mean)	*Organic Adopters* (Mean)	Diff	t-stats	p-value	*Offline-only* (Mean)
Panel A: Post-online-adoption Monthly Purchase Behaviors						
AvgSpendPerMonth	514.47	489.47	25.00	3.75	0.00	202.29
AvgQuantitiesPerMonth	4.88	4.11	0.77	7.58	0.00	1.85
AvgOrdersPerMonth	0.89	0.89	-0.00	-0.08	0.93	0.45
AvgUniqueItemsPerMonth	2.28	2.04	0.23	9.46	0.00	1.13
AvgUniqueBrandsPerMonth	1.78	1.63	0.15	9.18	0.00	0.91
AvgUniqueSubcategoriesPerMonth	1.52	1.39	0.13	10.35	0.00	0.80
AvgUniqueCategoriesPerMonth	1.27	1.16	0.11	10.82	0.00	0.68
AvgProfitPerMonth	175.95	157.42	18.53	8.18	0.00	76.73

the post-adoption period in Table A.6, complementing the pre-adoption statistics shown in Table 2.4 in §2.5.3. Finally, in Table A.7, we present more detailed summary statistics (e.g., median and quartiles) for all three groups during the post-online-adoption periods (i.e., between December 1st, 2019 and October 31, 2020). Note that the counts of *organic adopters* and *Black Friday adopters* from Table A.7 slightly differ from the raw count based on our cohorts' definition because 79 (2.58% of) *organic adopters* and 24 (1.74% of) *Black Friday adopters* churned before 2019-12-01, and therefore are not observed in the post-Black Friday panel. Including these customers does not affect our identification of post-online-adoption ATT, and the estimation results remain the same.

A.2.3 Event 3: Loyalty Program

In this section, we present the descriptive statistics for customers in the sample for the loyalty program analysis, supplementing those in §2.5.3. Table A.8 summarizes customer demographics. We also report customer-level summary statistics for average purchase variables

Table A.4. Customer-Level Summary Statistics for All Three Groups – *Covid adopters*, *Organic Adopters*, and *Offline-only* Customers for post-online-adoption period (i.e., between May 1st, 2020 and April 30th, 2023).

Variables	Mean	Std.	Min	25%	50%	75%	Max	Count
Covid adopters								
AvgSpendPerMonth	514.47	647.83	0.00	155.32	338.44	655.25	40153.38	32,313
AvgQuantitiesPerMonth	4.88	10.11	0.00	0.97	2.18	4.89	414.11	32,313
AvgOrdersPerMonth	0.89	0.81	0.00	0.36	0.69	1.17	17.81	32,313
AvgUniqueItemsPerMonth	2.28	2.41	0.00	0.78	1.57	2.97	58.62	32,313
AvgUniqueBrandsPerMonth	1.78	1.61	0.00	0.67	1.33	2.41	18.61	32,313
AvgUniqueSubcategoriesPerMonth	1.52	1.23	0.00	0.62	1.21	2.11	12.86	32,313
AvgUniqueCategoriesPerMonth	1.27	0.93	0.00	0.56	1.06	1.78	8.81	32,313
AvgProfitPerMonth	175.95	222.05	-95.36	54.56	116.51	224.52	15880.26	32,313
Organic Adopters								
AvgSpendPerMonth	489.47	582.43	0.00	141.99	317.95	629.33	10709.68	12,330
AvgQuantitiesPerMonth	4.11	8.10	0.00	0.83	1.89	4.22	245.86	12,330
AvgOrdersPerMonth	0.89	0.83	0.00	0.33	0.67	1.19	11.03	12,330
AvgUniqueItemsPerMonth	2.04	2.15	0.00	0.67	1.39	2.67	30.58	12,330
AvgUniqueBrandsPerMonth	1.63	1.51	0.00	0.58	1.19	2.19	16.31	12,330
AvgUniqueSubcategoriesPerMonth	1.39	1.16	0.00	0.54	1.08	1.92	9.47	12,330
AvgUniqueCategoriesPerMonth	1.16	0.88	0.00	0.50	0.94	1.61	6.81	12,330
AvgProfitPerMonth	157.42	191.05	-24.34	46.69	101.13	203.75	5028.99	12,330
Offline-only Customers								
AvgSpendPerMonth	202.29	330.96	0.00	35.98	99.68	242.19	24573.87	573,934
AvgQuantitiesPerMonth	1.85	4.59	0.00	0.30	0.77	1.85	487.17	573,934
AvgOrdersPerMonth	0.45	0.53	0.00	0.12	0.28	0.58	40.11	573,934
AvgUniqueItemsPerMonth	1.13	1.56	0.00	0.25	0.63	1.39	123.22	573,934
AvgUniqueBrandsPerMonth	0.91	1.09	0.00	0.23	0.55	1.17	48.44	573,934
AvgUniqueSubcategoriesPerMonth	0.80	0.87	0.00	0.21	0.50	1.06	19.90	573,934
AvgUniqueCategoriesPerMonth	0.68	0.70	0.00	0.19	0.44	0.94	10.00	573,934
AvgProfitPerMonth	76.73	120.02	-491.73	14.17	38.99	93.21	9834.37	573,934

Table A.5. Customer Demographics – Summary Statistics of Gender (Panel A) and Geographic State (Panel B).

Categories	*Black Friday Adopter* (%)	*Organic Adopter* (%)	*Offline-only* (%)
Panel A: Gender			
No Information	14.13	13.59	15.42
Female	59.42	57.89	46.46
Male	26.45	28.52	38.12
Total	100.00	100.00	100.00
Panel B: Geographic State			
São Paulo	57.97	51.91	64.83
Minas Gerais	8.48	13.10	4.69
Rio de Janeiro	7.90	6.99	4.50
Distrito Federal	4.78	4.64	4.38
Zip_State_Missing	4.93	3.76	9.90
Santa Catarina	2.54	3.40	2.60
Bahia	2.61	3.20	0.50
Goiás	2.83	3.17	2.19
Rio Grande do Sul	2.17	3.17	2.10
Espírito Santo	1.96	2.52	0.73
Paraná	1.74	1.99	2.39
Mato Grosso do Sul	1.52	1.83	1.10
Other States	0.58	0.33	0.09
Total	100.00	100.00	100.00

Table A.6. Customer-Level Summary Statistics of **Post-online-adoption** Purchase Behaviors between December 1st 2019 and October 31st 2020 (Panel A). Columns 1 - 2 describe statistics for *Black Friday adopters* and *Organic adopters*, respectively. Column 3 calculates the difference between *Black Friday adopters* and *Organic adopters*, and columns 4 and 5 show the two-sample t-test and p-value results.

Variable	*Black Friday Adopters* (Mean)	*Organic Adopters* (Mean)	Diff	t-stats	p-value	*Offline-only* (Mean)
Panel A: Post-online-adoption Monthly Purchase Behaviors						
AvgSpendPerMonth	417.53	468.23	-50.70	-2.94	0.00	195.21
AvgQuantitiesPerMonth	4.37	4.55	-0.17	-0.59	0.56	2.06
AvgOrdersPerMonth	0.88	0.97	-0.10	-3.36	0.00	0.48
AvgUniqueItemsPerMonth	2.28	2.39	-0.11	-1.33	0.18	1.24
AvgUniqueBrandsPerMonth	1.83	1.90	-0.08	-1.25	0.21	0.99
AvgUniqueSubcategoriesPerMonth	1.56	1.63	-0.07	-1.55	0.12	0.87
AvgUniqueCategoriesPerMonth	1.29	1.35	-0.06	-1.66	0.10	0.74
AvgProfitPerMonth	144.17	157.64	-13.47	-2.31	0.02	73.46

Table A.7. Customer-Level Summary Statistics for All Three Groups – *Black Friday adopters*, *Organic Adopters*, and *Offline-only* customers for post-online-adoption period (i.e., between December 1st, 2019 and October 31st, 2020).

Variables	Mean	Std.	Min	25%	50%	75%	Max	Count
				Black Friday adopters				
AvgSpendPerMonth	417.53	485.46	0.00	118.34	280.31	536.91	6029.03	1,356
AvgQuantitiesPerMonth	4.37	10.24	0.00	0.82	2.00	4.64	219.18	1,356
AvgOrdersPerMonth	0.88	0.82	0.00	0.27	0.64	1.18	6.91	1,356
AvgUniqueItemsPerMonth	2.28	2.51	0.00	0.64	1.50	3.09	20.27	1,356
AvgUniqueBrandsPerMonth	1.83	1.82	0.00	0.55	1.27	2.55	13.82	1,356
AvgUniqueSubcategoriesPerMonth	1.56	1.40	0.00	0.55	1.18	2.18	8.91	1,356
AvgUniqueCategoriesPerMonth	1.29	1.06	0.00	0.45	1.00	1.91	7.00	1,356
AvgProfitPerMonth	144.17	167.61	-56.75	40.64	96.10	184.10	2184.19	1,356
				Organic Adopters				
AvgSpendPerMonth	468.23	543.57	0.00	127.57	315.86	605.36	6608.86	2,982
AvgQuantitiesPerMonth	4.55	8.55	0.00	0.82	2.09	4.91	162.82	2,982
AvgOrdersPerMonth	0.97	0.93	0.00	0.36	0.73	1.27	7.82	2,982
AvgUniqueItemsPerMonth	2.39	2.66	0.00	0.67	1.64	3.09	27.00	2,982
AvgUniqueBrandsPerMonth	1.90	1.88	0.00	0.64	1.36	2.55	17.82	2,982
AvgUniqueSubcategoriesPerMonth	1.63	1.45	0.00	0.55	1.27	2.27	12.45	2,982
AvgUniqueCategoriesPerMonth	1.35	1.09	0.00	0.55	1.09	1.91	7.73	2,982
AvgProfitPerMonth	157.64	182.93	-105.01	42.74	105.73	203.51	2349.98	2,982
				Offline-only Customers				
AvgSpendPerMonth	195.21	330.14	0.00	14.31	86.21	243.99	16616.77	534,871
AvgQuantitiesPerMonth	2.06	5.51	0.00	0.18	0.73	2.09	522.36	534,871
AvgOrdersPerMonth	0.48	0.62	0.00	0.09	0.27	0.64	19.73	534,871
AvgUniqueItemsPerMonth	1.24	1.84	0.00	0.18	0.64	1.55	52.91	534,871
AvgUniqueBrandsPerMonth	0.99	1.31	0.00	0.12	0.55	1.36	23.18	534,871
AvgUniqueSubcategoriesPerMonth	0.87	1.06	0.00	0.09	0.55	1.25	16.09	534,871
AvgUniqueCategoriesPerMonth	0.74	0.85	0.00	0.09	0.45	1.09	9.82	534,871
AvgProfitPerMonth	73.46	120.07	-331.15	5.58	33.33	93.25	6313.29	534,871

during the post-adoption period in Table A.9, complementing the pre-adoption statistics shown in Table 2.4 in §2.5.3. Finally, in Table A.10, we present more detailed summary statistics (e.g., median and quartiles) for all three groups during the post-online-adoption periods (i.e., between October 1st, 2023 and March 31, 2024)

Table A.8. Customer Demographics – Summary Statistics of Gender (Panel A) and Geographic State (Panel B). Our data providers label around 45% of overall customers as no gender information in this study.

Categories	*Loyalty Program Adopter* (%)	*Organic Adopter* (%)	*Offline-only* (%)
Panel A: Gender			
No Information	50.16	48.30	42.85
Female	30.57	32.52	31.02
Male	19.27	19.17	26.13
Total	100.00	100.00	100.00
Panel B: Geographic State			
São Paulo	35.99	43.29	55.28
Zip_State_Missing	32.99	19.92	21.93
Minas Gerais	5.30	6.69	3.84
Rio de Janeiro	3.89	4.39	3.15
Santa Catarina	6.25	4.14	3.04
Distrito Federal	2.62	3.86	3.09
Espírito Santo	2.23	2.59	1.28
Ceará	2.04	2.37	0.83
Pernambuco	1.08	2.10	0.94
Bahia	1.21	1.99	1.12
Goiás	1.34	1.83	1.63
Mato Grosso do Sul	0.70	1.22	0.95
Other States	4.34	5.62	2.93
Total	100.00	100.00	100.00

Table A.9. Customer-Level Summary Statistics of **Post-online-adoption** Purchase Behaviors between October 1st, 2023 and March 31st, 2024 (Panel A). Columns 1 - 2 describe statistics for *Loyalty Program Adopters* and *Organic adopters*, respectively. Column 3 calculates the difference between *Loyalty Program Adopters* and *Organic adopters*, and columns 4 and 5 show the two-sample t-test and p-value results.

Variable	*Loyalty Program Adopters* (Mean)	*organic adopters* (Mean)	Diff	t-stats	p-value	*Offline-only* (Mean)
Panel A: Post-online-adoption Monthly Purchase Behaviors						
AvgSpendPerMonth	696.00	524.96	171.04	9.42	0.00	284.45
AvgQuantitiesPerMonth	5.22	3.58	1.64	7.70	0.00	2.18
AvgOrdersPerMonth	1.36	1.02	0.34	12.57	0.00	0.59
AvgUniqueItemsPerMonth	2.75	2.05	0.69	10.81	0.00	1.36
AvgUniqueBrandsPerMonth	2.16	1.65	0.51	11.62	0.00	1.09
AvgUniqueSubcategoriesPerMonth	1.84	1.42	0.41	12.33	0.00	0.96
AvgUniqueCategoriesPerMonth	1.54	1.22	0.32	12.27	0.00	0.83
AvgProfitPerMonth	223.51	180.19	43.32	6.86	0.00	110.25

Table A.10. Customer-Level Summary Statistics for All Three Groups – *Loyalty Program Adopters*, *Organic Adopters*, and *Offline-only Customers* for post-online-adoption period (i.e., between October 1st, 2023 and March 31st, 2024).

Variables	Mean	Std.	Min	25%	50%	75%	Max	Count
				Loyalty Program Adopters				
AvgSpendPerMonth	696.00	810.94	0.00	210.82	452.64	876.14	9712.41	1,567
AvgQuantitiesPerMonth	5.22	11.23	0.00	1.00	2.67	5.45	262.33	1,567
AvgOrdersPerMonth	1.36	1.31	0.00	0.50	1.00	1.83	20.50	1,567
AvgUniqueItemsPerMonth	2.75	2.89	0.00	1.00	2.00	3.67	43.33	1,567
AvgUniqueBrandsPerMonth	2.16	1.97	0.00	0.83	1.67	2.83	27.67	1,567
AvgUniqueSubcategoriesPerMonth	1.84	1.48	0.00	0.83	1.50	2.67	15.17	1,567
AvgUniqueCategoriesPerMonth	1.54	1.11	0.00	0.67	1.33	2.17	9.00	1,567
AvgProfitPerMonth	223.51	272.37	-44.82	66.43	144.49	280.82	3800.20	1,567
				Organic Adopters				
AvgSpendPerMonth	524.96	674.20	0.00	132.98	336.48	678.11	21482.62	16,746
AvgQuantitiesPerMonth	3.58	7.69	0.00	0.67	1.67	3.83	358.33	16,746
AvgOrdersPerMonth	1.02	0.99	0.00	0.33	0.83	1.33	15.17	16,746
AvgUniqueItemsPerMonth	2.05	2.38	0.00	0.67	1.33	2.67	41.67	16,746
AvgUniqueBrandsPerMonth	1.65	1.64	0.00	0.50	1.17	2.17	22.33	16,746
AvgUniqueSubcategoriesPerMonth	1.42	1.25	0.00	0.50	1.17	2.00	14.33	16,746
AvgUniqueCategoriesPerMonth	1.22	0.97	0.00	0.50	1.00	1.67	9.00	16,746
AvgProfitPerMonth	180.19	235.65	-1078.75	45.07	114.10	233.52	7693.62	16,746
				Offline-only Customers				
AvgSpendPerMonth	284.45	484.92	0.00	36.70	146.34	356.93	37203.48	396,396
AvgQuantitiesPerMonth	2.18	5.39	0.00	0.33	1.00	2.17	859.33	396,396
AvgOrdersPerMonth	0.59	0.69	0.00	0.17	0.40	0.83	50.67	396,396
AvgUniqueItemsPerMonth	1.36	1.90	0.00	0.33	0.83	1.80	114.83	396,396
AvgUniqueBrandsPerMonth	1.09	1.29	0.00	0.25	0.67	1.50	56.00	396,396
AvgUniqueSubcategoriesPerMonth	0.96	1.02	0.00	0.20	0.67	1.33	22.67	396,396
AvgUniqueCategoriesPerMonth	0.83	0.82	0.00	0.17	0.67	1.17	10.67	396,396
AvgProfitPerMonth	110.25	180.70	-690.21	14.02	56.76	138.84	11025.70	396,396

A.3 Appendix to DID results for Adopters vs. Offline-only customers

A.3.1 Loyalty Program Control Variables

Below, we provide the definition of time-varying control variables related to customers' participation in the loyalty program. These variables ($\boldsymbol{x}_{it}$) are included in the DID and event study analysis for the loyalty program study.

- $isActivated_{it}$: 1 if customer i activates the offer during month t and 0 otherwise. Use to compare within-activators, how much spend differs between months with/without activation.
- $isFirstActivation_{it}$: A binary variable equal to 1 if customer i activates the offer for the first time in month t, and 0 otherwise. This variable isolates the novelty (first-time) in the month of initial activation, relative to subsequent activation months.
- $Months\_Since\_First\_Activation_{it}$: A non-negative integer indicating the number of months elapsed since customer i's first activation of the offer, as of month t; equals 0 for months prior to the first activation. For example, if the first activation occurs in August 2023, then

  $Months\_Since\_First\_Activation_{i,September\_2023} = 1$, and so on.
- $\%Month\_WithActivation_{it}$: The proportion of months in which customer i has activated the offer up to (but not including) month t, calculated as the total number of activation months up to t divided by the number of months since first activation (excluding t itself). For the first month of activation, this variable is defined as 0 (since the denominator is 0); all first-month effects are captured by $isFirstActivation_{it}$. Starting from the second month post-activation, this variable reflects the historical activation rate prior to month t.

A.3.2 Standard DiD Estimation Results – Adopters vs. Offline-only customers

In the main text, we only considered the TWFE model. We now present the estimation results for the standard two-by-two DID model specified below:

$$y_{it} = \alpha_0 + \alpha_1 \mathbb{I}\{Adopter_i\} + \alpha_2 \mathbb{I}\{Post_t\} + \alpha_3 \mathbb{I}\{Adopter_i\} \times \mathbb{I}\{Post_t\} + \boldsymbol{\alpha}_4' \boldsymbol{x}_{it} + \epsilon_{it} \quad \text{(A.1)}$$

Table A.11 and A.12 are estimation results for the standard two-by-two DID model for spend and profit between *adopters* vs. *offline-only* customers, respectively.

Table A.11. DID Results — Spend (*Adopters* vs. *Offline-only*) (Standard 2×2 DID)

Study:	Covid	Black Friday	Loyalty Program
DV:		Spend	
Model:	(1)	(2)	(3)
Variables			
Constant	280.1 (0.557)	291.7 (0.586)	332.6 (0.800)
	[0.000]	[0.000]	[0.000]
Adopter	177.2 (2.885)	177.8 (9.177)	123.4 (4.42)
	[0.000]	[0.000]	[0.000]
Post	-62.90 (0.496)	-96.95 (0.426)	-51.36 (0.632)
	[0.000]	[0.000]	[0.000]
Adopter × Post	140.4 (2.843)	81.29 (7.543)	10.02 (5.191)
	[0.000]	[0.000]	[0.053]
Loyalty program controls			Yes
Fit statistics			
Observations	22,204,608	9,505,999	4,403,449
R^2	0.01905	0.00980	0.00918

Clustered (Customer) standard errors in parentheses; p-values in brackets.

Table A.12. DID Results — Profit (*Adopters* vs. *Offline-only*) (Standard 2×2 DID)

Study:	Covid	Black Friday	Loyalty Program
DV:		Profit	
Model:	(1)	(2)	(3)
Variables			
Constant	92.57 (0.177)	96.13 (0.187)	125.8 (0.288)
	[0.000]	[0.000]	[0.000]
Adopter	61.35 (0.925)	60.94 (3.041)	41.57 (1.607)
	[0.000]	[0.000]	[0.000]
Post	-10.56 (0.174)	-22.83 (0.152)	-16.81 (0.238)
	[0.000]	[0.000]	[0.000]
Adopter × Post	35.87 (0.974)	19.61 (2.547)	-7.231 (1.820)
	[0.000]	[0.000]	[0.000]
Loyalty program controls			Yes
Fit statistics			
Observations	22,204,608	9,505,999	4,403,449
R^2	0.01228	0.00450	0.00651

Clustered (Customer) standard errors in parentheses; p-values in brackets.

A.4 Appendix to DID Results for *Event Adopters* vs. *Organic Adopters*

A.4.1 Appendix to DID results for spend

We now run the standard two-by-two DID as follows:

$$y_{it} = \delta_0 + \delta_1 \mathbb{1}\{Event\_Adopter_i\} + \delta_2 \mathbb{1}\{Post_t\} + \delta_3 \mathbb{1}\{Event\_Adopter_i\} \times \mathbb{1}\{Post_t\} + \boldsymbol{\delta_4}' \boldsymbol{x}_{it} + \epsilon_{it} \tag{A.2}$$

Table A.13 shows the estimation results for the standard two-by-two DID model for spend between *event Adopters* vs. *organic Adopters.*

Discussion of Forward buying

Now, we discuss how to set up the event study regression for testing the forward buying under Black Friday and loyalty program analysis. For the Black Friday analysis, we define

Table A.13. DID Main Analysis – Spend (*Event Adopters* vs. *Organic Adopters*) (Standard 2×2 DID)

Study:	Covid	Black Friday	Loyalty Program
DV:		Spend	
Model:	(1)	(2)	(3)
Variables			
Organic_Adopter	432.8 (5.325)	459.1 (10.48)	444.7 (4.415)
	[0.000]	[0.000]	[0.000]
Event_Adopter	32.84 (6.281)	32.01 (20.72)	131.7 (19.08)
	[0.000]	[0.122]	[0.000]
Post	85.85 (5.34)	10.54 (9.334)	-29.92 (5.041)
	[0.000]	[0.257]	[0.000]
Event_Adopter × Post	-10.64 (6.271)	-82.25 (15.47)	-254.7 (20.82)
	[0.089]	[0.000]	[0.000]
Loyalty program controls			Yes
Fit statistics			
Observations	1,783,949	76,842	197,658
R^2	0.00131	0.00081	0.05635

Clustered (Customer) standard errors in parentheses; p-values in brackets.

event time 0 as November 2019, given that Black Friday occurred on 2019-11-29. The pre-adoption period is from 2019-01-01 to 2019-10-31 (event time -10 to -1), and the post-adoption period is from 2019-12-01 to 2020-10-31 (event time 1 to 11). By estimating the event study regression in Equation (2.3), we present the estimation result in Table A.14 and show that *Black Friday* adopters spend significantly higher than *organic adopters* during the adoption month, but spend significantly lower during the post-adoption periods, which indicates a pattern of forward buying.

Regarding the loyalty program analysis, for both *Loyalty Program adopters* and *organic adopters*, the month of adoption is labeled as event time 0. For example, event time 0 is June 2023 for *organic adopters* who adopted online shopping in that month, and September 2023 for *Loyalty Program adopters* who adopted online shopping in that month.

We compare each group's spending at event time 0 to their most recent pre-adoption outcome, specifically those observed in May 2023. This period serves as a common baseline for pre-treatment behaviors across all adopter groups. As defined in §2.5.2.3, the post-adoption periods span from October 2023 and March 2024, and we label October 2023 as the event time 1 for both *organic* and *Loyalty Program adopters* and label all the following months until March 2024 as event time 6. The event study coefficient at time 1 captures the difference in the monthly spend at October 2023 between two groups relative to the pre-adoption difference in the reference period May 2023.

We first compare the spend of *organic* and *Loyalty Program adopters* during their adoption month relative to the pre-adoption reference period (i.e., spend in May 2023), and Table A.15 shows that *Loyalty Program adopters* tend to spend 177.7 MCU more than *organic adopters* during the adoption month. Next, we notice that both *Loyalty Program adopters* and *organic adopters* can activate the monthly loyalty offer during the post-adoption period. To test the post-adoption differences in purchase behavior, we estimate the event study regression in Equation (2.3) with control variables for the customers' activations. Table A.16 shows that *Loyalty Program adopters* spend significantly less than *organic adopters* in most of the post-adoption periods.

Table A.14. Event Study for Spend – *Black Friday Adopters* vs. *Organic Adopters*

DV:	Spend
Model:	(1)
Variables	
BlackFriday_Adopter × Lag –10	–39.54 (50.73)
	[0.436]
BlackFriday_Adopter × Lag –9	–44.26 (38.91)
	[0.255]
BlackFriday_Adopter × Lag –8	–26.81 (35.61)
	[0.452]
BlackFriday_Adopter × Lag –7	–25.93 (30.61)
	[0.397]
BlackFriday_Adopter × Lag –6	5.206 (34.45)
	[0.880]
BlackFriday_Adopter × Lag –5	–4.364 (27.55)
	[0.874]
BlackFriday_Adopter × Lag –4	–18.65 (27.25)
	[0.494]
BlackFriday_Adopter × Lag –3	–9.695 (25.51)
	[0.704]
BlackFriday_Adopter × Lag –2	–38.55 (23.94)
	[0.107]
BlackFriday_Adopter × Adoption Month	128.6 (32.46)
	[0.000]
BlackFriday_Adopter × Lead 1	–134.3 (25.36)
	[0.000]
BlackFriday_Adopter × Lead 2	–114.4 (25.65)
	[0.000]
BlackFriday_Adopter × Lead 3	–105.5 (25.58)
	[0.000]
BlackFriday_Adopter × Lead 4	–88.21 (28.63)
	[0.002]
BlackFriday_Adopter × Lead 5	–55.82 (28.28)
	[0.048]
BlackFriday_Adopter × Lead 6	–126.9 (27.95)
	[0.000]
BlackFriday_Adopter × Lead 7	–88.91 (28.21)
	[0.002]
BlackFriday_Adopter × Lead 8	–69.56 (27.24)
	[0.011]
BlackFriday_Adopter × Lead 9	–25.79 (28.13)
	[0.359]
BlackFriday_Adopter × Lead 10	–90.03 (28.42)
	[0.002]
BlackFriday_Adopter × Lead 11	–67.00 (28.53)
	[0.019]
Fixed-effects	
Customer	Yes
YearMonth	Yes
Fit statistics	
Observations	81,283
R^2	0.42086
Within R^2	0.00211

Clustered (Customer) standard-errors are presented in standard brackets, while p-values are shown in square brackets.

Table A.15. Event Study for Spend during the Adoption Month - *Loyalty Program Adopters* vs. *Organic Adopters*

DV:	Spend
Model:	(1)
Variables	
Loyalty_Adopter × Lag -5	-11.43 (26.52)
	[0.672]
Loyalty_Adopter × Lag -4	-29.64 (25.37)
	[0.245]
Loyalty_Adopter × Lag -3	0.7290 (24.62)
	[0.976]
Loyalty_Adopter × Lag -2	-25.12 (26.40)
	[0.340]
Loyalty_Adopter × Adoption Month	177.7 (31.54)
	[0.000]
Fixed-effects	
Customer	Yes
Event_Time	Yes
Fit statistics	
Observations	109,857
R^2	0.53391
Within R^2	0.00111

Clustered (Customer) standard-errors are presented in standard brackets, while p-values are shown in square brackets.

Table A.16. Event Study for Spend during the Post-Online-Adoption Periods – *Loyalty Program Adopters* vs. *Organic Adopters*

DV:	Spend
Model:	(1)
Variables	
Loyalty_Adopter × Lag -5	-11.45 (26.52)
	[0.670]
Loyalty_Adopter × Lag -4	-29.64 (25.37)
	[0.245]
Loyalty_Adopter × Lag -3	0.7526 (24.62)
	[0.975]
Loyalty_Adopter × Lag -2	-25.10 (26.40)
	[0.340]
Loyalty_Adopter × Lead 1	-102.5 (29.56)
	[0.001]
Loyalty_Adopter × Lead 2	-34.48 (28.49)
	[0.224]
Loyalty_Adopter × Lead 3	-45.26 (31.24)
	[0.144]
Loyalty_Adopter × Lead 4	-73.56 (30.45)
	[0.015]
Loyalty_Adopter × Lead 5	-69.94 (33.10)
	[0.037]
Loyalty_Adopter × Lead 6	-75.00 (35.29)
	[0.032]
Loyalty program controls	Yes
Fixed-effects	
Customer	Yes
YearMonth	Yes
Fit statistics	
Observations	197,658
R^2	0.47130
Within R^2	0.02699

Clustered (Customer) standard-errors are presented in standard brackets, while p-values are shown in square brackets.

Next, we run the event study regression for offline spend, and Figure A.1 presents the event study figure for the offline spend of Black Friday and loyalty program study.

Finally, we follow a similar approach described above to test for *covid adopters*. Figure A.2 shows that there is no forward buying for *covid adopters* relative to *organic adopters*.

Figure A.1. Event Study for Offline Spend

(a) Black Friday (b) Loyalty Program

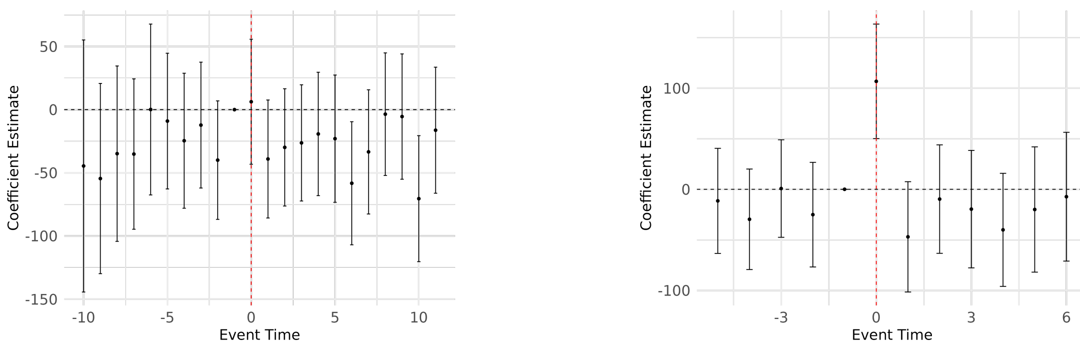
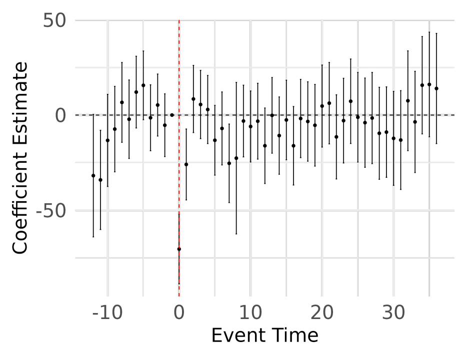


Figure A.2. Event Study for Spend – *Covid Adopters* vs. *Organic Adopters* during the month of adoption



A.4.2 Appendix to DID results for the Share of Offline Spend

Table A.17 presents the estimation results for the linear trend model by comparing a type of *event adopters* with *organic adopters*. Table A.18 presents the changes in share of offline spend by different age groups, showing senior adopters tend to have a slower rate of shifting their spend to the online channel.

Table A.17. Differential Trends in the Share of Offline Spend (Estimation Results for Equation (A.3)) – *Event Adopters* vs. *Organic Adopters*

DV:	Share of Offline Spend		
Model:	Covid	Black Friday	Loyalty Program
Variables			
Organic_Adopter	0.4833 (0.0084)	0.5930 (0.0285)	0.5662 (0.0093)
	[0.000]	[0.000]	[0.000]
Event_Adopter	0.0807 (0.0119)	0.1424 (0.0404)	0.0306 (0.0132)
	[0.000]	[0.000]	[0.020]
t	-0.0020 (0.0004)	-0.0108 (0.0048)	-0.0020 (0.0031)
	[0.000]	[0.025]	[0.518]
Event_Adopter × t	0.0012 (0.0006)	-0.0109 (0.0068)	0.0013 (0.0043)
	[0.046]	[0.109]	[0.763]
Fit statistics			
Observations	72	22	12
R^2	0.81922	0.69953	0.72620
Adjusted R^2	0.81124	0.64945	0.62353

Standard errors in parentheses; p-values in square brackets.

Table A.18. Differential Trend in the Share of Offline Spend by Age – *Covid Adopters* vs. *Organic Adopters*

DV:	Share of Offline Spend	
Model:	(1) Senior Adopters	(2) Young Adopters
Variables		
Organic_Adopter	0.4947 (0.0092)	0.4728 (0.0081)
	[0.000]	[0.000]
Covid_Adopter	0.0706 (0.0130)	0.0882 (0.0114)
	[0.000]	[0.000]
t	-0.0019 (0.0005)	-0.0022 (0.0004)
	[0.001]	[0.000]
Covid_Adopter × t	0.0020 (0.0006)	0.0004 (0.0006)
	[0.002]	[0.523]
Fit statistics		
Observations	72	72
R^2	0.79964	0.82429
Adjusted R^2	0.79080	0.81654

Standard errors are shown in parentheses; p-values are shown in square brackets.

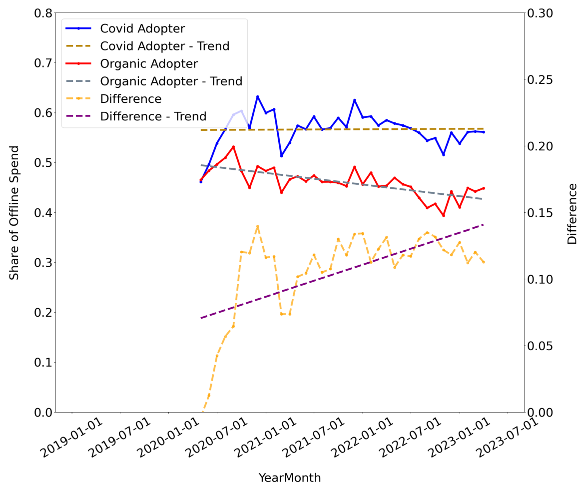
In §2.6.2.2, we find that the share of offline spending is persistently higher for *covid adopters* compared to *organic adopters* in the post-adoption period, and the downward trend is steeper for *organic adopters*, indicating that *covid adopters* shift their spending to online channels slower than *organic adopters*. To formally test whether these differential trends in share of offline spend are systematic, we estimate the following model with linear time trend on the post-adoption monthly panel data for both groups:

$$y_{gt} = \zeta + \rho_0 t + \rho_1 \mathbb{I}\{Event\_Adopter_g\} + \rho_2 \mathbb{I}\{Event\_Adopter_g\} \times t + \epsilon_{gt}. \tag{A.3}$$

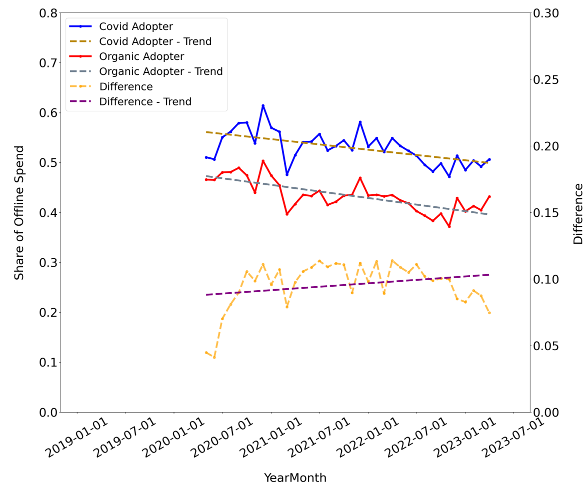
Here, y_{gt} is the fraction of the offline spend relative to the total spend (ranging from 0 to 1) for group g in period t. t is the numerical time index, and we set the first post-adoption month of each study as $t = 0$. The results for covid, Black Friday and loyalty program studies from this regression are shown in Table A.17. The estimate of the difference in trend (ρ_2) between covid and organic adoption cohorts is 0.0012 (or 0.12%). This finding suggests that although the different reasons (i.e., organic vs. exogenous shock) for adopting online shopping do not seem to have much impact on *how much* consumers spend with the firm in total, they do cause differences in *where* they spend it. Figure A.3 plots the fraction of offline spending in the post-treatment period separately for older and younger adopters.

Figure A.3. Differential Trend Visualization by Age – *Covid Adopters* vs. *Organic Adopters*

(a) Senior Adopters (Age $\geq$ 41)



(b) Young Adopters (Age $<$ 41)



A.4.3 Appendix to DID results for Profit

Table A.19 shows the estimation result for the standard two-by-two DID model for the profit between *event Adopters* vs. *organic Adopters*.

Table A.19. DID Main Analysis – Profit (*Event Adopters* vs. *Organic Adopters*) (Standard 2×2 DID)

Study:	Covid	Black Friday	Loyalty Program
DV:		Profit	
Model:	(1)	(2)	(3)
Variables			
Organic_Adopter	145.7 (1.698)	154.7 (3.519)	164.5 (1.627)
	[0.000]	[0.000]	[0.000]
Event_Adopter	10.96 (2.007)	7.338 (6.79)	34.13 (6.413)
	[0.000]	[0.28]	[0.000]
Post	20.27 (1.788)	3.315 (3.163)	-21.05 (1.788)
	[0.000]	[0.295]	[0.000]
Event_Adopter × Post	7.248 (2.117)	-20.52 (5.228)	-86.83 (7.352)
	[0.000]	[0.000]	[0.000]
Loyalty program controls			Yes
Fit statistics			
Observations	1,783,949	76,842	197,658
R^2	0.00134	0.00039	0.04568

Clustered (Customer) standard errors in parentheses; p-values in brackets.

A.5 Appendix to HonestDiD and Linear Trend Estimation

A.5.1 Parallel Trends for Black Friday Analysis

Figure A.4a and A.4d visualize the monthly average total spend and profit values for *Black Friday adopters*, *organic adopters* and *offline-only* customers. Then, we provide a pre-trend test for two main metrics (i.e., spend and profit) between *adopters* (including both *Black Friday adopters* and *organic adopters*) relative to *offline-only* customers. In total, we have 10 pre-treatment periods (i.e., between January 2019 and October 2019), labeled as -10 to -1, and 11 post-adoption periods (i.e., between December 2019 and October 2020), labeled as 1 to 11. The base period is October 2019, and the event time 0 refers to the adoption month (i.e., November 2019). Figure A.4b and A.4e exhibit a minor violation during the early stage of the pre-adoption period, but most pre-adoption coefficients are significant. To show the robustness of our main finding in the presence of violation of parallel pre-trend, we show the bounded ATT in §2.6.3 and confirm that the incremental spend and profit are still positive and significant in the presence of violation of parallel trend.

Finally, we provide a pre-trend test between *Black Friday adopters* and *organic adopters*. Figure A.4c and A.4f presents the event study regression estimates for the parallel pre-trend of the Black Friday study.

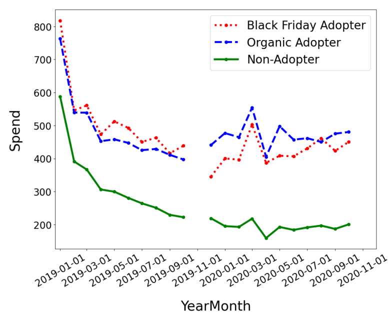
A.5.2 Parallel Trends for Loyalty Program Analysis

Figure A.5a and A.5d visualize the monthly average total spend and profit values for *Loyalty Program adopters*, *organic adopters* and *offline-only* customers.

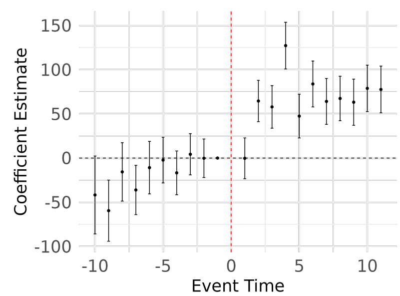
Next, we provide a pre-trend test for two main metrics (i.e., spend and profit) between *adopters* (including both *Loyalty Program adopters* and *Organic adopters*) relative to *offline-only* customers and present the event study estimates (see Figure A.5b and A.5e for total spend and profit, respectively). We exhibit a violation of the parallel trend assumption. To show the robustness of our main finding in the presence of violation of parallel pre-trend, we show the bounded ATT in §2.6.3 and confirm that the incremental spend is still positive

Figure A.4. Parallel Trend Assessment – Black Friday. Panels (A.4a, A.4d) plot monthly outcomes, while panels (A.4b, A.4e) and (A.4c, A.4f) report event study estimates (based on Equation (2.3)) comparing (i) *adopters* (both *Black Friday* and *organic*) to *offline-only customers* and (ii) *Black Friday* to *organic adopters*, respectively.

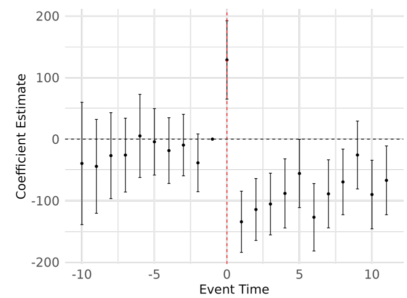
(a) Visualization – Spend



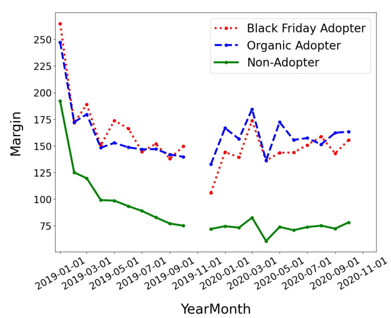
(b) *Adopters* vs. *Offline-only* – Spend



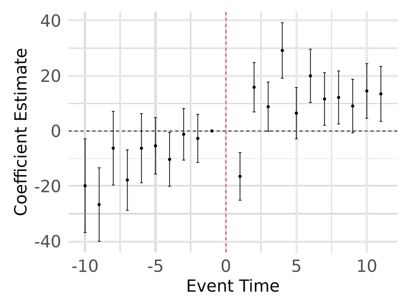
(c) *Black Friday* vs. *Organic Adopters* – Spend



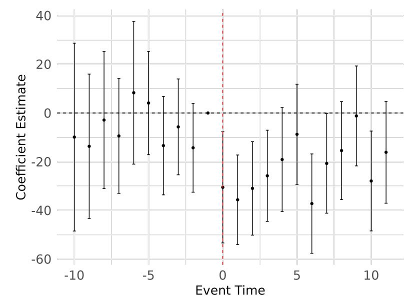
(d) Visualization – Profit



(e) *Adopters* vs. *Offline-only* – Profit



(f) *Black Friday* vs. *Organic Adopters* – Profit

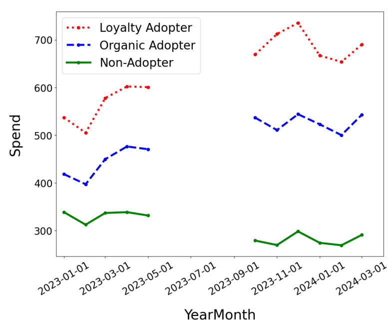


and significant in the presence of violation of parallel trend.

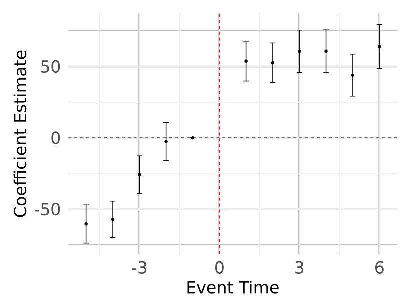
Finally, Figure A.5c and A.5f present the event study regression estimates for the parallel pre-trend of the loyalty program online activation study.

Figure A.5. Parallel Trend Assessment – Loyalty Program. Panels (A.5a, A.5d) plot monthly outcomes, while panels (A.5b, A.5e) and (A.5c, A.5f) report event study estimates (based on Equation (2.3)) comparing (i) *adopters* (both *Loyalty Program* and *organic*) to *offline-only customers* and (ii) *Loyalty Program* to *organic adopters*, respectively.

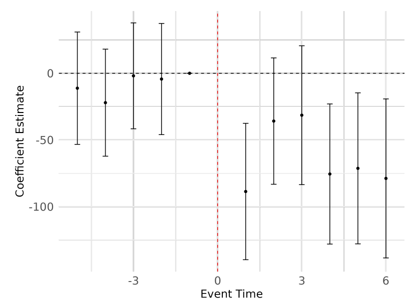
(a) Visualization – Spend



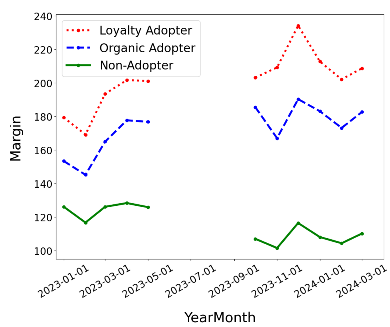
(b) *Adopters* vs. *Offline-only* – Spend



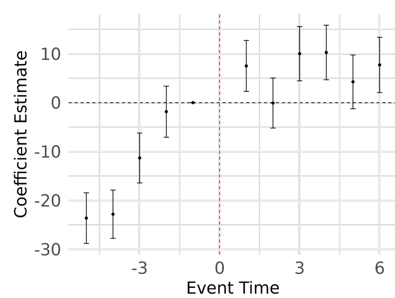
(c) *Loyalty Program* vs. *Organic Adopters* – Spend



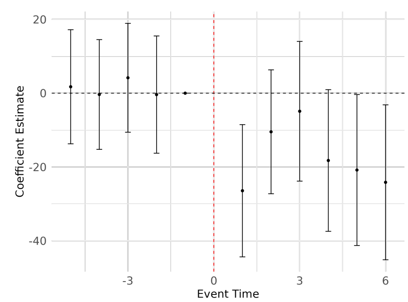
(d) Visualization – Profit



(e) *Adopters* vs. *Offline-only* – Profit



(f) *Loyalty Program* vs. *Organic Adopters* – Profit



A.5.3 Linear Trend Model Estimates

To quantify the magnitude of the differential pre-trends, we estimate a linear trend model using the entire pre-treatment data to test for any differential time trends in the average monthly spend and profit between the treatment and control customers before the adoption.

$$y_{it} = \gamma_i + \eta_t + \phi_{e,v}\mathbb{I}\{Treatment\}_i \times t + \epsilon_{it} \tag{A.4}$$

Here, $\mathbb{I}\{Treatment\}_i$ refers to the treatment customers. When we compare *adopters* with *offline-only customers*, $\mathbb{I}\{Treatment\}_i = 1$ is customer i is an adopter. When we compare *covid adopters* with *organic adopters*, $\mathbb{I}\{Treatment\}_i = 1$ if customer i is a *covid adopter*. t is the year-month number for the pre-adoption period, and $\phi_{e,v}$ is the linear trend coefficient for event e variable v. If no linear trend exists during the pre-adoption period, the coefficient $\phi_{e,v}$ should be statistically insignificant. The results are shown in Table A.20. We see some significant trend coefficients, except for one insignificant coefficient for pre-adoption monthly profit between adopters (Black Friday and organic) and offline-only customers (see column 4 in Table A.20).

Table A.20. Linear Trend Tests Across Adoption Pathways

	Adopters vs. Offline-only						**Within Adopters**	
	Covid		**Black Friday**		**Loyalty Program**		**Covid vs. Organic**	
DVs:	Spend	Profit	Spend	Profit	Spend	Profit	Spend	Profit
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Adopter × t	4.170	1.128	4.244	1.081	17.51	6.811	4.393	1.415
	(0.377)	(0.141)	(1.571)	(0.6015)	(1.504)	(0.578)	(0.8605)	(0.3183)
	[0.000]	[0.000]	[0.007]	[0.072]	[0.000]	[0.000]	[0.000]	[0.000]
Fixed-effects								
Customer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	5,146,359	5,146,359	3,855,908	3,855,908	2,073,524	2,073,524	378,387	378,387
R^2	0.48244	0.39923	0.48754	0.40597	0.55397	0.51055	0.51305	0.42830
Within R^2	6.28×10^{-5}	3.05×10^{-5}	5.24×10^{-6}	2.25×10^{-6}	0.00012	0.00012	0.00012	7.71×10^{-5}

Clustered (Customer) standard errors are reported in parentheses; p-values are reported in square brackets.
Columns (1)–(6) compare adopters in each pathway group to offline-only customers.
Columns (7)–(8) compare Covid adopters directly to Organic adopters.

A.6 Appendix to Managerial Implications

Table A.21 presents an additional DID result for profit, which are inputs for the discussion of managerial implications. Instead of estimating Equation (2.1), with a single indicator $\mathbb{I}\{Adopter_i\}$, we now include two separate dummies – $\mathbb{I}\{Event\_Adopter_i\}$ and $\mathbb{I}\{Organic\_Adopter_i\}$ – each interacted with the post-adoption indicator. This allows us to estimate the ATT of online adoption on profit for *organic adopters* and *event adopters*, relative to *offline-only* customers.

Table A.21. DID Results (All Three Groups) – Profit

Study:	Covid	Black Friday	Loyalty Program
Dependent Var.:		Profit	
Model:	(1)	(2)	(3)
Variables			
Organic_Adopter × Post	30.28 (1.903)	23.57 (3.160)	18.55 (1.509)
	[0.000]	[0.000]	[0.000]
Event_Adopter × Post	39.78 (1.272)	7.813 (4.014)	2.986 (6.798)
	[0.000]	[0.051]	[0.292]
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	22,204,608	9,505,999	4,403,449
R^2	0.31903	0.34750	0.44546
Within R^2	0.00047	2.88×10^{-5}	0.00194

Clustered (Customer) standard errors in parentheses; p-values in brackets.

Additional Analysis for ROI/breakeven analyses for promotion:

We also perform a similar calculation for the Loyalty Program, with the cost per customer to set up the loyalty program and to convince customers to activate it online is 50 MCU and 150 MCU. Using our DID estimates in Appendix §A.6, the average monthly incremental profit from online adoption is 18.55 MCU for *organic adopters* and 2.986 MCU for *Loyalty Program adopters*.

When the cost per customer is 50 MCU, if managers naively used the incremental profit from *organic adopters*, they would compute a breakeven time of 2.70 months (i.e., 50/18.55), whereas the correct breakeven period using *Loyalty Program adopters*' incremental profit is 16.74 months (i.e., 50/2.986) – more than six times longer. Therefore, managers should appropriately calibrate their expectations of increased consumer spending when running promotions that encourage the adoption of online shopping.

When the cost per customer is 150 MCU, if managers naively used the incremental profit from *organic adopters*, they would compute a breakeven time of 8.09 months (i.e., 150/18.55), whereas the correct breakeven period using *Loyalty Program adopters*' incremental profit is 50.23 months (i.e., 150/2.986) – more than six times longer. Therefore, managers should appropriately calibrate their expectations for increased consumer spending when running promotions that encourage online shopping adoption.

A.7 Appendix to Validity and Robustness Checks

A.7.1 Appendix to Effect Size Measurement

Figure A.6 presents the bootstrapped distribution of true ATT for *covid adopters* relative to *organic adopters*. Figure A.7 presents the bootstrapped distribution of true ATT for *Black Friday adopters* relative to *organic adopters*.

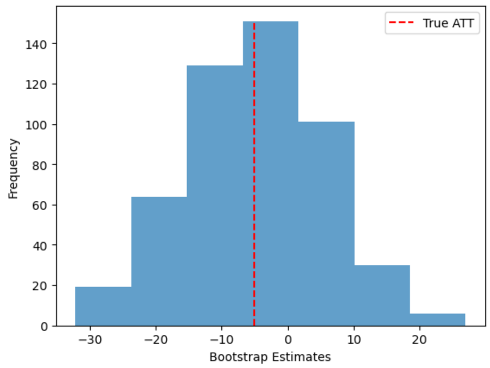
A.7.2 Appendix to DID Models with Propensity Score Matching and Inverse Propensity Score Weighting – Organic Adopters vs. Event Adopters

In this section, we provide a set of robustness checks under PSM and IPTW. To conduct this analysis, we label the *event adopters* as 1 and *organic adopters* as 0 and model the propensity score of customer i being an *event adopter*. The propensity score model is specified as follows:

$$PropensityScore_i = f(X_i^T \theta), \tag{A.5}$$

Figure A.6. True ATT – *Covid Adopters* vs. *Organic Adopters*

(a) Spend



(b) Profit

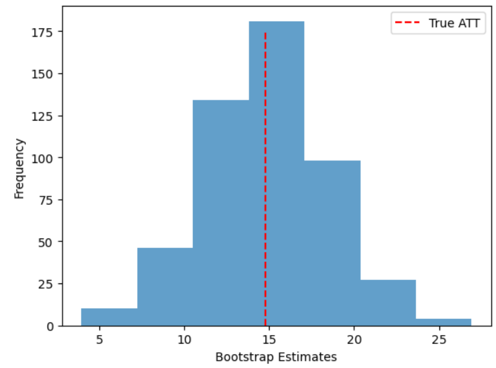
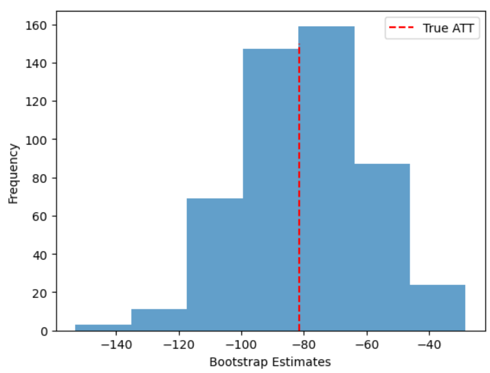
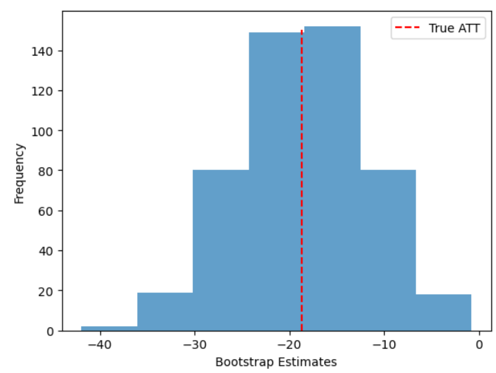


Figure A.7. True ATT – *Black Friday Adopters* vs. *Organic Adopters*

(a) Spend



(b) Profit



where $f(\cdot)$ is the logit link function where $f(X_i^T\theta) = \frac{1}{1+\exp(-X_i^T\theta)}$. $X_i \in \mathcal{R}^{M\times 1}$ is a column vector with M covariates of customer i. These covariates include $Gender_i$, Age_i, $Tenure(Days)_i$, $AverageSpendPerMonth_i$, etc. $\theta \in \mathcal{R}^{M\times 1}$ is a column vector of coefficients. Table A.22 in Appendix A.7.2.1 presents the propensity score model estimation for covid, Black Friday, and loyalty program analysis, respectively. Figure A.8 presents the distribution of propensity scores before matching, and we observe some differences between *event adopters* and *organic adopters*. After matching, Figure A.9 suggests that *event adopters* and *organic adopters* become more overlapped.

Next, we use these propensity scores to augment our DID analysis in two ways.

- First, we perform a customer-level propensity score matching and then re-run the DID analysis. Matching based on propensity score is much more efficient than matching based on a set of covariates (Rosenbaum and Rubin, 1985). Based on the predicted propensity score of each customer i, we use *MatchIt* package in R (Stuart et al., 2011) to perform nearest-neighbor matching for each *event adopter*. We then run the DID analysis on the newly matched sample (see Table A.24, A.25, and A.26 in Appendix A.7.2.2). We find that the results from this exercise are consistent with the main findings.
- Next, we use a slightly different approach to control for any potential selection issues. Specifically, we use an inverse propensity of treatment weight-adjusted (IPTW) DID regression. Here, the sample is the same as the one from the main analysis, but each observation is inversely weighted by its propensity score. We then perform a two-sample t-test on all the pre-treatment variables based on the weighted sample and confirm that the two samples are statistically insignificant on all the variables (see Table A.23 for the weighted sample of *organic adopters* and *event adopters*, respectively in the Appendix A.7.2.3). The DID results of *event adopters* and *organic adopters* from this analysis are shown in Table A.27, A.28 and A.29 for covid, Black Friday, and Loyalty Program analysis, respectively in Appendix A.7.2.3. Again, we find that the results and conclusions are quite similar to those in the main section.

A.7.2.1 Propensity Score Model Estimation

Table A.22 presents the propensity score model estimates for *covid adopters*, *Black Friday adopters*, and *Loyalty Program adopters* versus *organic adopters*. Figure A.8 presents the distribution of propensity scores of *event adopters* and *organic adopters* under three studies. After performing nearest-neighbor matching, A.9 shows that the distribution of propensity scores between *event adopters* and *organic adopters* becomes more overlapped.

Figure A.8. Propensity Score Model Predictions - Before Matching

(a) Covid (b) Black Friday (c) Loyalty Program

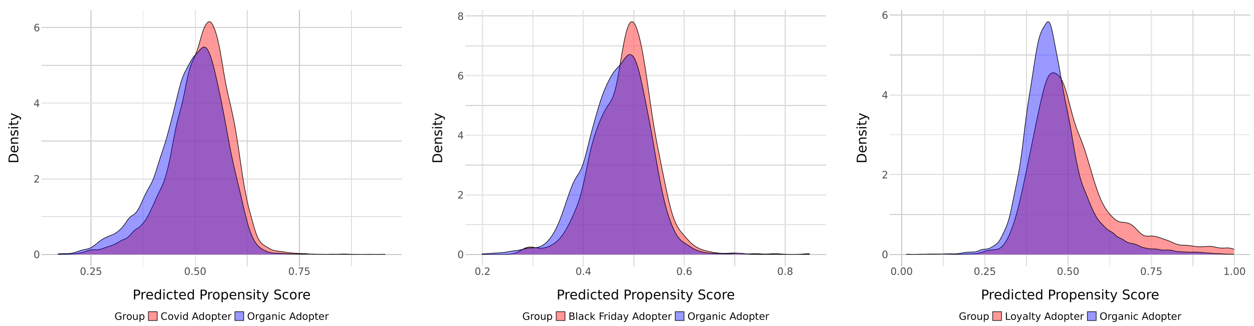
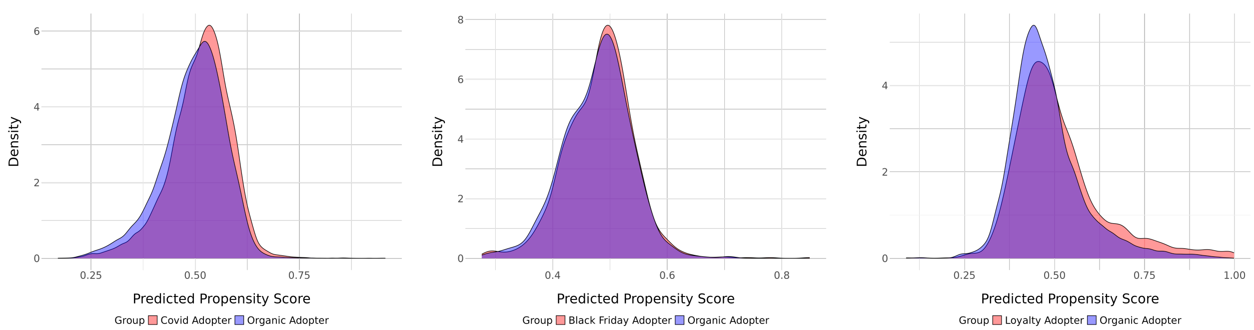


Figure A.9. Propensity Score Model Predictions - After Matching

(a) Covid (b) Black Friday (c) Loyalty Program



A.7.2.2 DID Results with PSM

Table A.24, A.25 and A.26 present the DID estimates of Equation (2.2) of key purchase metrics based on the matched sample under covid, Black Friday, and loyalty program analysis. All results are consistent with the main findings.

Table A.22. Propensity Score Logit Model – *Event Adopters* vs. *Organic Adopters*

	DVs:		
	Covid_Adopter	BlackFriday_Adopter	LoyaltyProgram_Adopter
Gender_NoInformation	-0.044 (0.027)	0.095 (0.088)	-0.031 (0.042)
	[0.103]	[0.278]	[0.459]
Gender_Female	0.052 (0.018)	0.098 (0.062)	-0.050 (0.035)
	[0.004]	[0.114]	[0.152]
LogTenureDays	0.243 (0.009)	0.175 (0.031)	-0.087 (0.026)
	[0.000]	[0.000]	[0.001]
AvgSpendPerMonth	-0.0001 (0.0001)	0.0003 (0.0002)	0.002 (0.0001)
	[0.317]	[0.133]	[0.000]
AvgQuantityPerMonth	0.010 (0.002)	0.008 (0.006)	0.0001 (0.003)
	[0.000]	[0.183]	[0.974]
AvgUniqueOrdersPerMonth	-0.147 (0.018)	-0.121 (0.063)	0.174 (0.028)
	[0.000]	[0.055]	[0.000]
AvgUniqueItemsPerMonth	0.036 (0.013)	0.022 (0.043)	-0.066 (0.021)
	[0.006]	[0.610]	[0.002]
AvgUniqueBrandsPerMonth	-0.086 (0.022)	-0.031 (0.070)	-0.0001 (0.042)
	[0.000]	[0.659]	[0.998]
AvgUniqueSubcategoriesPerMonth	0.003 (0.025)	0.081 (0.080)	0.315 (0.055)
	[0.905]	[0.311]	[0.000]
AvgUniqueCategoriesPerMonth	0.088 (0.026)	-0.040 (0.089)	-0.096 (0.059)
	[0.001]	[0.655]	[0.103]
AvgUniqueVisitedStoresPerMonth	0.289 (0.041)	0.284 (0.137)	-0.074 (0.067)
	[0.000]	[0.038]	[0.272]
AvgProfitPerMonth	0.0002 (0.0001)	-0.001 (0.0005)	-0.006 (0.0003)
	[0.045]	[0.045]	[0.000]
Constant	-1.063 (0.410)	-1.191 (0.203)	-0.079 (0.189)
	[0.010]	[0.000]	[0.672]
Age	Yes	Yes	Yes
Household Income	Yes	Yes	Yes
Observations	46,045	4,441	18,313
Log Likelihood	-48,790.450	-3,983.461	-19,285.020
Akaike Inf. Crit.	97,622.890	8,006.922	38,610.040

Standard-errors in standard brackets; p-values in square brackets.

Note: *AvgUniqueVisitedStoresPerMonth* measures the average number of distinct stores a customer visits per month during the pre-adoption period under each study.

Table A.23. Customer-level Summary Statistics of Pre-Adoption Variables for *Event Adopters* and *Organic Adopters*: Re-weighted Sample Based on Inverse Propensity Score. Columns 2 and 3 show the weighted mean of variables of *event adopters* and *organic adopters*, with weights being the inverse of propensity.

Variable	*Event Adopters* mean	*Organic adopters* mean	Diff	T-stats	P-value
Panel A: Covid Analysis					
LogTenure(Days)	5.12	5.12	-0.00	-0.40	0.69
AvgSpendPerMonth	459.22	461.68	-2.47	-0.42	0.67
AvgQuantitiesPerMonth	4.72	4.91	-0.19	-1.67	0.09
AvgOrdersPerMonth	0.95	0.95	0.00	0.04	0.97
AvgUniqueItemsPerMonth	2.90	2.90	0.00	0.01	1.00
AvgUniqueBrandsPerMonth	2.32	2.33	-0.00	-0.04	0.97
AvgUniqueSubcategoriesPerMonth	2.03	2.03	-0.00	-0.10	0.92
AvgUniqueCategoriesPerMonth	1.65	1.65	-0.00	-0.09	0.93
AvgProfitPerMonth	162.25	162.72	-0.47	-0.22	0.83
Panel B: Black Friday Analysis					
LogTenure(Days)	5.01	5.01	0.00	0.02	0.99
AvgSpendPerMonth	477.58	477.41	0.17	0.01	0.99
AvgQuantitiesPerMonth	4.75	4.70	0.05	0.21	0.83
AvgOrdersPerMonth	1.00	1.00	0.00	0.01	1.00
AvgUniqueItemsPerMonth	3.00	3.00	0.00	0.01	0.99
AvgUniqueBrandsPerMonth	2.42	2.42	-0.00	-0.00	1.00
AvgUniqueSubcategoriesPerMonth	2.10	2.10	-0.00	-0.01	0.99
AvgUniqueCategoriesPerMonth	1.70	1.70	-0.00	-0.00	1.00
AvgProfitPerMonth	166.97	167.08	-0.11	-0.02	0.99
Panel C: Loyalty Program Analysis					
LogTenure(Days)	6.45	6.45	0.00	0.14	0.89
AvgSpendPerMonth	514.74	535.10	-20.36	-1.38	0.17
AvgQuantitiesPerMonth	3.73	3.85	-0.12	-0.70	0.48
AvgOrdersPerMonth	0.93	0.94	-0.01	-0.42	0.67
AvgUniqueItemsPerMonth	2.28	2.29	-0.02	-0.27	0.78
AvgUniqueBrandsPerMonth	1.80	1.81	-0.01	-0.37	0.71
AvgUniqueSubcategoriesPerMonth	1.54	1.55	-0.01	-0.39	0.70
AvgUniqueCategoriesPerMonth	1.30	1.31	-0.01	-0.38	0.70
AvgProfitPerMonth	180.94	184.54	-3.60	-0.70	0.49

Table A.24. *Covid Adopters* vs. *Organic Adopters* – PSM + DID

DVs:	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
Covid_Adopter × Post	-1.842 (6.910)	38.08 (5.813)	0.0693 (0.0046)	9.258 (2.423)
	[0.790]	[0.000]	[0.000]	[0.000]
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,689,232	1,689,232	901,177	1,689,232
R^2	0.39081	0.36431	0.50579	0.29503
Within R^2	1.72e-07	0.00012	0.00125	2.59e-05

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Table A.25. *Black Friday Adopters* vs. *Organic Adopters* – PSM + DID

DVs:	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
BlackFriday_Adopter × Post	-78.26 (17.94)	-16.68 (17.37)	0.0959 (0.0159)	-22.46 (6.431)
	[0.000]	[0.337]	[0.000]	[0.001]
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	42,329	42,329	23,221	42,329
R^2	0.41012	0.41414	0.52090	0.34764
Within R^2	0.00099	5.74e-05	0.00671	0.00058

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Table A.26. *Loyalty Program Adopters* vs. *Organic Adopters* – PSM + DID

DVs:	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
Loyalty_Adopter × Post	-52.62 (20.60)	-8.541 (19.41)	0.0013 (0.0117)	-16.32 (7.228)
	[0.011]	[0.660]	[0.912]	[0.024]
Loyalty program controls	Yes	Yes	Yes	Yes
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	159,333	159,333	90,778	159,333
R^2	0.47539	0.44214	0.56883	0.44853
Within R^2	0.02722	0.01915	0.00922	0.02278

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

A.7.2.3 DID Results with IPTW

Table A.27, A.28 and A.29 present the DID estimates of Equation (2.2) of key purchase metrics with inverse propensity score weighting under covid, Black Friday, and loyalty program analysis. All results are consistent with the main findings.

Table A.27. *Covid Adopters* vs. *Organic Adopters* - IPTW + DID

DVs	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
Covid_Adopter × Post	-6.305 (6.550)	37.96 (5.465)	0.0651 (0.0043)	6.651 (2.307)
	[0.315]	[0.000]	[0.000]	[0.008]
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,783,949	1,783,949	949,448	1,783,949
R^2	0.38696	0.36110	0.50768	0.29041
Within R^2	2.31×10^{-6}	0.00014	0.00125	1.53×10^{-5}

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Table A.28. *Black Friday Adopters* vs. *Organic Adopters* - IPTW + DID

DVs:	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
BlackFriday_Adopter × Post	-68.89 (14.30)	-4.429 (14.12)	0.0848 (0.0122)	-20.64 (5.231)
	[0.000]	[0.765]	[0.000]	[0.000]
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	76,842	76,842	42,299	76,842
R^2	0.40586	0.39667	0.53426	0.34948
Within R^2	0.00067	3.58×10^{-6}	0.00455	0.00042

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Table A.29. *Loyalty Program Adopters* vs. *Organic Adopters* - IPTW + DID

DVs	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
Loyalty_Adopter × Post	-52.53 (19.43)	-11.43 (18.60)	0.0017 (0.0114)	-15.22 (6.835)
	[0.009]	[0.546]	[0.888]	[0.024]
Loyalty program controls	Yes	Yes	Yes	Yes
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	197,658	197,658	110,339	197,658
R^2	0.47155	0.43591	0.57046	0.44366
Within R^2	0.02694	0.01910	0.00903	0.02249

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

A.7.2.4 Doubly Robust DID for Covid Analysis

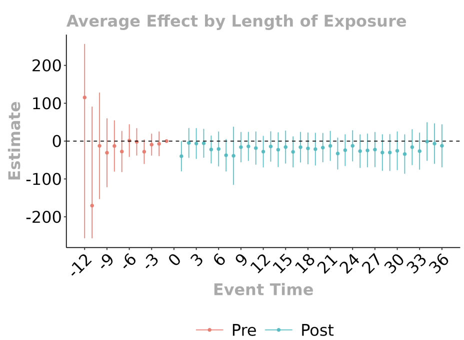
Given the potential concern of parallel trend assumption under COVID-19 analysis, we seek an alternative identification assumption and conduct a doubly robust DID for the covid analysis. We now show that the main conclusions remain consistent by using the doubly robust DID (Sant'Anna and Zhao, 2020). We provide the ATT estimates for each metric based on the package *did* in R provided by Callaway and Sant'Anna (2021). Table A.30 presents the doubly robust DID estimates for all metrics. Figure A.10 below presents the event study figure under doubly robust DID results for main metrics (i.e., spend and profit). We can see that the parallel pre-trends are also well satisfied, conditional on customer demographics and pre-adoption purchase variables. The estimation results are consistent with the main findings.

Table A.30. Doubly Robust DID – *Covid Adopters* vs. *Organic Adopters*

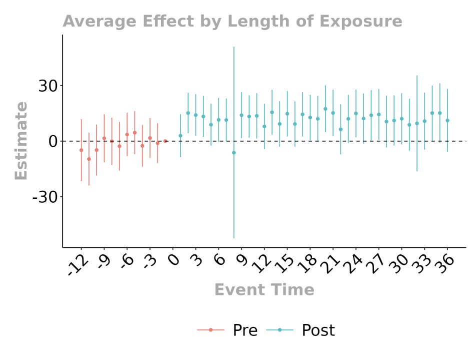
DVs	ATT	SE	T-stat	P-value	CI-	CI+
Spend	-5.963	7.814	-0.763	0.445	-21.279	9.353
Share of Offline Spend	0.083	0.004	20.550	0.000	0.075	0.091
Profit	11.623	2.942	3.951	0.000	5.856	17.390

Figure A.10. Doubly Robust Event Study for the Main Metrics – *Covid Adopters* vs. *Organic Adopters*

(a) Spend



(b) Profit



A.7.3 Appendix to DID Models with Propensity Score Matching and Inverse Propensity Score Weighting – Adopters vs. Offline-only Customers

In this section, we label the *adopters* (both *event adopters* and *organic adopters*) as 1 and *offline-only* customers as 0 and model the propensity score of customer i being an *adopter*. The propensity score model is specified as follows:

$$PropensityScore_i = f(X_i^T \theta), \tag{A.6}$$

where $f(\cdot)$ is the logit link function where $f(X_i^T \theta) = \frac{1}{1+\exp(-X_i^T \theta)}$. $X_i \in \mathcal{R}^{M \times 1}$ is a column vector with M covariates of customer i. These covariates include $Gender_i$, Age_i, $Tenure(Days)_i$, $AverageSpendPerMonth_i$, etc. $\theta \in \mathcal{R}^{M \times 1}$ is a column vector of coefficients. Table A.31 in Appendix A.7.3 presents the propensity score model estimation for COVID, Black Friday, and loyalty program analysis, respectively. Figure A.11 presents the distribution of propensity scores before matching, and we observe some differences between *adopters* and *offline-only* customers. After matching, Figure A.12 suggests that *adopters* and *offline-only* customers become more overlapped.

Next, we use these propensity scores of being adopters to augment our DID analysis in two ways.

- First, we perform a customer-level propensity score matching and then re-run the DID analysis. Matching based on propensity score is much more efficient than matching based on a set of covariates (Rosenbaum and Rubin, 1985). Based on the predicted propensity score of each customer i, we use *MatchIt* package in R (Stuart et al., 2011) to perform nearest-neighbor matching for each *adopter*. We then run the DID analysis on the newly matched sample and find that the results from this exercise are consistent with the main findings (see Table A.32).
- Next, we use a slightly different approach to control for any potential selection issues. Specifically, we use an inverse propensity of treatment weight-adjusted (IPTW) DID regression. Here, the sample is the same as the one from the main analysis, but each observation is inversely weighted by its propensity score. The DID results of *adopters*

and *offline-only* customers from this analysis are shown in Table A.33. Again, we find that the results and conclusions are quite similar to those in the main section.

A.7.3.1 Propensity Score Model Estimation

Table A.31 presents the propensity score model estimates for being an *adopter* under three studies. Figure A.11 presents the distribution of propensity scores of being an *adopter* under three studies. After performing nearest-neighbor matching, A.12 shows that the distribution of propensity scores of being an *adopter* becomes more overlapped.

Figure A.11. Propensity Score Model Predictions - Before Matching

(a) Covid (b) Black Friday (c) Loyalty Program

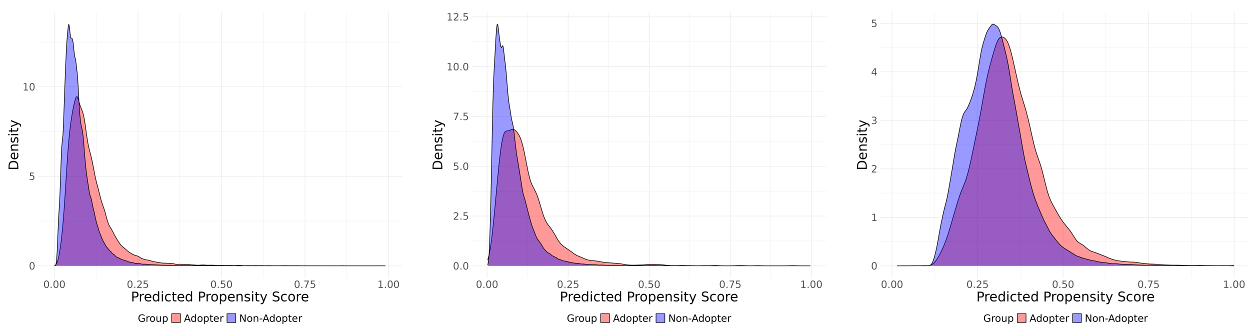


Figure A.12. Propensity Score Model Predictions - After Matching

(a) Covid (b) Black Friday (c) Loyalty Program

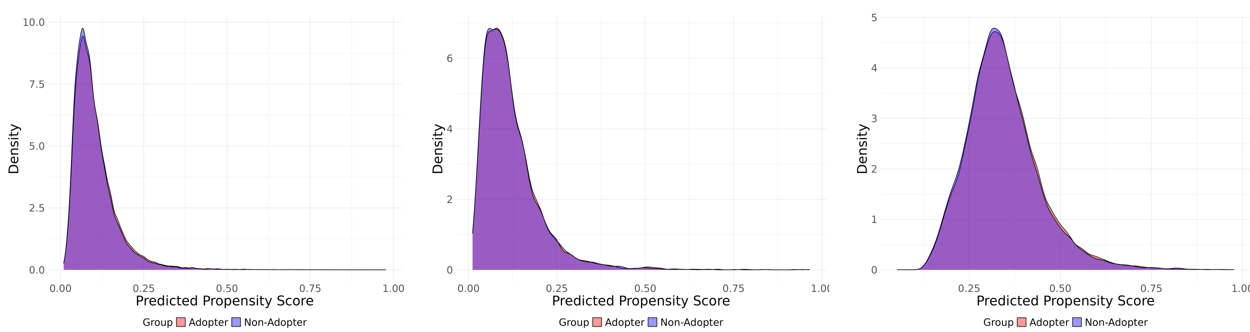


Table A.31. Propensity Score Logit Model – Adopters vs. Non-Adopters

	DV:		
	Covid	Black Friday	Loyalty Program
Gender_NoInformation	-0.003 (0.006)	0.123 (0.016)	0.290 (0.010)
	[0.617]	[0.000]	[0.000]
Gender_Female	0.494 (0.004)	0.544 (0.012)	0.355 (0.008)
	[0.000]	[0.000]	[0.000]
LogTenureDays	0.051 (0.002)	-0.039 (0.006)	-0.323 (0.006)
	[0.000]	[0.000]	[0.000]
AvgSpendPerMonth	0.001 (0.00002)	0.001 (0.00003)	0.001 (0.00003)
	[0.000]	[0.000]	[0.000]
AvgQuantityPerMonth	0.009 (0.001)	-0.003 (0.001)	-0.014 (0.001)
	[0.000]	[0.003]	[0.000]
AvgUniqueOrdersPerMonth	-0.290 (0.005)	-0.179 (0.012)	-0.031 (0.008)
	[0.000]	[0.000]	[0.000]
AvgUniqueItemsPerMonth	-0.050 (0.004)	-0.059 (0.008)	-0.023 (0.006)
	[0.000]	[0.000]	[0.000]
AvgUniqueBrandsPerMonth	0.085 (0.007)	0.166 (0.014)	-0.024 (0.011)
	[0.000]	[0.000]	[0.029]
AvgUniqueSubcategoriesPerMonth	-0.055 (0.008)	-0.052 (0.016)	0.157 (0.015)
	[0.000]	[0.001]	[0.000]
AvgUniqueCategoriesPerMonth	0.253 (0.008)	0.138 (0.017)	0.123 (0.015)
	[0.000]	[0.000]	[0.000]
AvgUniqueVisitedStoresPerMonth	0.430 (0.012)	0.605 (0.026)	0.159 (0.017)
	[0.000]	[0.000]	[0.000]
AvgProfitPerMonth	0.0002 (0.00005)	-0.001 (0.0001)	-0.003 (0.0001)
	[0.000]	[0.000]	[0.000]
Constant	-2.464 (0.079)	-5.123 (0.321)	1.166 (0.045)
	[0.000]	[0.000]	[0.000]
Age	Yes	Yes	Yes
HouseholdIncome	Yes	Yes	Yes
Observations	619,979	539,905	414,709
Log Likelihood	-720,390.800	-145,457.300	-349,382.200
Akaike Inf. Crit.	1,440,824.000	290,956.700	698,804.300

Standard errors in parentheses; p-values in square brackets.

Note: *AvgUniqueVisitedStoresPerMonth* measures the average number of distinct stores a customer visits per month during the pre-adoption period under each study.

A.7.3.2 DID Results with PSM/IPTW

Table A.32 and Table A.33 present the DID estimates from two-way fixed effects regression based on the matched samples and with inverse propensity weighting, respectively. We see the findings are consistent with the main results.

Table A.32. PSM + DID Estimates of Adopters vs. Non-Adopters

	Covid		Black Friday		Loyalty Program	
Variable	Spend	Profit	Spend	Profit	Spend	Profit
Adopter × Post	191.8 (3.701)	53.37 (1.335)	140.4 (10.03)	38.22 (3.553)	127.5 (5.336)	32.38 (1.933)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Loyalty program controls					Yes	Yes
Fixed-effects						
Customer	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	3,310,659	3,310,659	153,086	153,086	384,412	384,412
R^2	0.4107	0.3262	0.4228	0.3681	0.4948	0.4656
Within R^2	0.00351	0.00169	0.00344	0.00180	0.02170	0.01570

Clustered (Customer) standard-errors are presented in parentheses, while p-values are shown in square brackets.

Table A.33. IPTW + DID Estimates of Adopters vs. Non-Adopters

	Covid		Black Friday		Loyalty Program	
Variable	Spend	Profit	Spend	Profit	Spend	Profit
Adopter × Post	200.0 (2.758)	55.99 (0.9605)	135.7 (7.784)	36.64 (2.443)	144.2 (5.584)	36.62 (1.795)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Loyalty program controls					Yes	Yes
Fixed-effects						
Customer	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	22,204,608	22,204,608	9,505,999	9,505,999	4,403,449	4,403,449
R^2	0.4067	0.3394	0.4442	0.3625	0.6956	0.6298
Within R^2	0.00165	0.00083	0.00160	0.00078	0.00419	0.00308

Clustered (Customer) standard-errors are presented in parentheses, while p-values are shown in square brackets.

A.7.4 Appendix to Alternate definitions for the organic adopter and covid adopter cohorts

A.7.4.1 Appendix to Covid Analysis - Re-defining the Organic Adopter Cohort

In this section, we provide a robustness check by redefining *organic adopters* as those customers who were offline-only before 2019-10-31 and made their first online purchase between 2019-11-01 to 2020-02-29. The definition of *covid adopters* remains the same. Under this new definition, we continue to have 33,246 *covid adopters* and a larger number of 25,958 *organic adopters*. As before, we exclude the periods used for defining these two cohorts (2019-11-01 to 2020-04-30) and re-run the same set of DID models as in the main analysis. Table A.34 presents the DID results between *covid adopters* and *organic adopters*. We find that all the results are consistent with the main analysis.

Table A.34. *Covid Adopters vs. Organic Adopters* – Redefining Organic Adopter Group

DVs:	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
Covid_Adopter × Post	10.03 (5.944)	44.76 (4.773)	0.0580 (0.0034)	13.20 (2.082)
	[0.091]	[0.000]	[0.000]	[0.000]
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	2,178,690	2,178,690	1,153,061	2,178,690
R^2	0.39039	0.36518	0.50935	0.30924
Within R^2	5.66e-06	0.00021	0.00108	6.31e-05

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

A.7.4.2 Appendix to Covid Analysis - Re-defining the Covid Adopter Cohort

In this section, we provide an alternative definition of *covid adopters* cohort for the covid analysis. We provide a robustness check by redefining *covid adopters* as those customers who were offline-only before 2019-12-31 and made their first online purchase between 2020-05-01 and 2020-06-30 (i.e., two months later than the exact onset of COVID-19). The definition of *organic adopters* remains the same as the definition in §2.5.2.1. Under this definition, we continue to have 12,799 *organic adopters* but now have 27,448 *covid adopters*. As before, we exclude the periods used for defining these two cohorts (i.e., 2020-01-01 to 2020-06-30), re-run the same set of DID models as in the main analysis, and present the results in Table A.35. We find that all results are consistent with the main analysis.

Table A.35. *Covid Adopters* vs. *Organic Adopters* – Re-defining *Covid Adopters* Group

DVs:	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
Covid_Adopter × Post	7.16 (6.69)	70.2 (5.50)	0.078 (0.004)	9.36 (2.31)
	[0.286]	[0.000]	[0.000]	[0.000]
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,461,440	1,461,440	744,927	1,461,440
R^2	0.37338	0.36248	0.51561	0.27316
Within R^2	3.38e-06	0.00055	0.00202	3.41e-05

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

A.7.4.3 Appendix to Covid Analysis - Comparing Covid Adopters (March/April 2020) vs. Later Covid Adopters (May/June 2020)

Table A.36 compares *covid adopters* (i.e., adopt during March/April 2020) and *later covid adopters* (i.e., adopt during May/June 2020). Although the *later covid adopters* spend more than the *covid adopters* by 9.96 MCU, it is not large in magnitude with a p-value of 0.048. So we conclude there is no major difference in post-adoption spend, share of offline spend (i.e., channel utilization), and profit.

Table A.36. *Covid Adopters* (March/April) vs. *Later Covid Adopters* (May/June)

DVs:	Spend	Offline Spend	Share of Offline Spend	Profit
Model:	(1)	(2)	(3)	(4)
Variables				
Later_Covid_Adopter × Post	9.96 (5.040)	30.4 (4.270)	-0.0002 (0.003)	0.070 (1.770)
	[0.048]	[0.000]	[0.947]	[0.968]
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	2,240,950	2,240,950	1,170,421	2,240,950
R^2	0.39921	0.36673	0.50555	0.35185
Within R^2	8.12e-06	0.00011	9.28e-09	3.00e-09

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

A.7.5 Appendix to Using Natural Log Scale Dependent Variables

In this section, we provide a robustness check for using natural log transformed scale of the dependent variables (i.e., ln(dependent variable + 1)), and conclude that the direction of effects are consistent with the main finding. Note that the profit variable can take both positive and negative values, so we do not consider natural log transformation of profit. First, we present the result for spend between *adopters* and *offline-only customers* in Table A.37, and show that the results are consistent with Table 2.5.

Next, Table A.38 presents DID results for the total spend of comparing *event adopters* and *organic adopters* by three events. Note that the profit variable can take both positive and negative values, so we do not consider natural log transformation of profit. We show that the results in Table A.38 are consistent with Table 2.7.

Table A.37. DID Analysis - Spend (in Natural Log Scale) (*Adopters* vs. *Offline-only*)

Study:	Covid	Black Friday	Loyalty Program
DV		ln(Spend + 1)	
Model:	(1)	(2)	(3)
Variables			
Adopter × Post	0.6883 (0.0104)	0.5742 (0.0302)	0.6099 (0.0190)
	[0.000]	[0.000]	[0.000]
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	22,204,608	9,505,999	4,403,449
R^2	0.30744	0.36627	0.29925
Within R^2	0.00087	0.00010	0.00208

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Table A.38. DID Analysis – Spend (in Natural Log Scale) (*Event Adopters* vs. *Organic Adopters*)

Study:	Covid	Black Friday	Loyalty Program
DV		ln(Spend + 1)	
Model:	(1)	(2)	(3)
Variables			
Event_Adopter × Post	0.008 (0.023)	-0.268 (0.062)	-0.151 (0.064)
	[0.728]	[0.000]	[0.017]
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	1,783,949	76,842	197,658
R^2	0.28677	0.32002	0.31155
Within R^2	2.16×10^{-7}	0.00049	0.02144

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

A.7.6 Appendix to Handling Zero-Inflated Dependent Variables

In this section, we outline the procedure and results of the two-part hurdle model and the estimation results. The two-part hurdle model includes two separate estimation steps:

- Step 1: Estimate a linear probability model with outcome variable $\mathbb{I}\{Spend_{it}\}$ with all independent variables specified in Equation (2.1) when comparing *adopters* vs. *offline-only*, and independent variables specified in Equation (2.2) when comparing *event adopters* vs. *organic adopters*;
- Step 2: Estimate Equation (2.1) conditional on customers making non-zero spend when comparing *adopters* vs. *offline-only*, and independent variables specified in Equation (2.2) when comparing *event adopters* vs. *organic adopters*.

Next, we present the estimation results for *adopters* vs. *offline-only* customers and for each *event Adopters* vs. *organic adopters*.

Adopters vs. Offline-only

Table A.39 presents the estimation result of linear probability model of $\mathbb{I}\{Spend_{it}\}$ between *adopters* and *offline-only* customers. Table A.40 presents the DID estimation results conditional on customer-month with non-zero spend between *adopters* and *offline-only* customers.

Event Adopters vs. Organic Adopters

Table A.41 presents the estimation result of linear probability model of $\mathbb{I}\{Spend_{it}\}$ between *event adopters* and *organic adopters*. Table A.42 presents the DID estimation results conditional on customer-month with non-zero spend between *event adopters* and *organic adopters*.

Table A.39. Linear Probability Model with DID Specifications – *Adopters* vs. *Offline-only*

Study:	Covid	Black Friday	Loyalty Program
DV		$\mathbb{I}\{Spend_{it}\}$	
Model:	(1)	(2)	(3)
Variables			
Adopter × Post	0.1033 (0.0016)	0.0910 (0.0046)	0.0933 (0.0029)
	[0.000]	[0.000]	[0.000]
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	22,204,608	9,505,999	4,403,449
R^2	0.28000	0.34101	0.26334
Within R^2	0.00073	9.3×10^{-5}	0.00156

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Table A.40. DID Estimates for Spend (Conditional on Purchase) – *Adopters* vs. *Offline-only*

Study:	Covid	Black Friday	Loyalty Program
DV		Spend	
Model:	(1)	(2)	(3)
Variables			
Adopter × Post	135.4 (3.953)	65.40 (9.385)	50.28 (6.334)
	[0.000]	[0.000]	[0.000]
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	8,187,536	3,744,951	1,939,763
R^2	0.47201	0.50834	0.62202
Within R^2	0.00084	4.37×10^{-5}	0.00150

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Table A.41. Linear Probability Model with DID Specifications – *Event Adopters* vs. *Organic Adopters*

Study:	Covid	Black Friday	Loyalty Program
DVs:		$\mathbb{I}\{Spend_{it}\}$	
Model:	(1)	(2)	(3)
Variables			
Event_Adopter × Post	0.0039 (0.0034)	-0.0361 (0.0095)	-0.0202 (0.0095)
	[0.251]	[0.000]	[0.032]
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	1,783,949	76,842	197,658
R^2	0.25272	0.28908	0.27038
Within R^2	2.36×10^{-6}	0.00036	0.01574

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Table A.42. DID Estimates for Spend (Conditional on Purchase) – *Event Adopters* vs. *Organic Adopters*

Study:	Covid	Black Friday	Loyalty Program
DVs:		Spend	
Model:	(1)	(2)	(3)
Variables			
Event_Adopter × Post	-3.982 (8.373)	-64.19 (19.68)	-36.03 (23.92)
	[0.623]	[0.001]	[0.126]
Loyalty program controls			Yes
Fixed-effects			
Customer	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes
Fit statistics			
Observations	949,448	42,299	110,339
R^2	0.44326	0.45643	0.56886
Within R^2	7.04×10^{-7}	0.00047	0.01280

Clustered (Customer) standard-errors in parentheses; p-values in brackets.

Appendix B
APPENDIX TO CHAPTER 3

B.1 Appendix to Setting and Data

B.1.1 Details of Dataset Construction

B.1.1.1 Selection of Brands and Stock Keeping Units (SKUs)

For each product category (i.e., dry dog food, dry cat food, dog hygiene and cat hygiene), we focus on a single pack size that accounts for the largest sales revenue and offers sufficient brand representation. Our goal is to study consumer brand choice and price elasticity, so consistent with previous brand choice literature (Allenby and Rossi, 1998; Kim et al., 2002; Chintagunta et al., 2005; Dubé et al., 2010) we focus on a market (i.e., single most popular pack size within each category) that initially (as of January 1st, 2019 - June 30th, 2019) features at least three brands. We exclude any market with fewer than three brands or with one brand contributing more than 60% of total sales. Applying these criteria yields the top-selling pack size for each category during the observation period: 15KG adult daily dry dog food, 10.1KG adult daily dry cat food, and an 80cm × 60cm dog hygiene mat containing 30 pieces.

However, we observe high market concentration for certain cat hygiene pack sizes. The 12KG, 4KG, and 3KG sizes are among the top-selling segments (e.g., ranked 1st, 3rd, and 4th in offline sales). However, each has less than three brands, and a single brand contributes over 60% of sales revenue (85%, 63%, and 88%, respectively), making these pack sizes unsuitable for brand choice analysis. Furthermore, the excluded pack sizes differ significantly in sand types, compositions, and functional characteristics, limiting their substitutability in the consumer's mind. For example, 3KG pack sizes SKUs are primarily made from cereal-and vegetable-based materials that are eco-friendly, biodegradable, lightweight, and prone to scattering, whereas 4KG and 12KG pack sizes consist of granulated smectite clay, which is heavy, non-biodegradable and derived from mined materials. Hence, these pack sizes cater to distinct functional needs and consumer preferences. [1] The differences in product attributes and target consumer preferences complicate defining a cohesive market that includes these pack sizes. In addition, even if 4KG and 12KG pack sizes are grouped into a single market, the top brand still accounts for 77% of the sales, leaving little room for brand choice analysis.

Consequently, we focus on silica cat sand in the 1.6KG, 1.8KG, and 2.0KG, which together account for 29% of the store's cat sand sales. Among these, the 1.8KG size is the second largest in the offline cat sand sales. Despite slight variations in pack sizes across brands, all three pack sizes offer the same sand category (i.e., silica), and we observe significant overlap in consumer purchases. Specifically, among customers who purchased the 1.8KG size at least once during the sample period, 28% also purchased the 1.6KG pack size, and 20% also purchased the 2.0KG pack size during the observation period, indicating a clear overlap of consumer preferences in this segment.

After defining the market with the top one selling pack size for each category, we rank brands by offline sales in descending order and include those that collectively account for at least 80% of the sales, following earlier literature (Gordon et al., 2013). The remaining brands are grouped into a single 'Other Brand' for each product category, ensuring broad market coverage.

B.1.1.2 Price and Cost Index Calculations

We construct brand-level price indices for each category by aggregating SKU-level price data from *offline* transactions (Gordon et al., 2013; Hitsch et al., 2021). First, we construct a store-SKU-month level price matrix from the available transactions data. To begin with, this price matrix could contain both outliers and missing values. Any observed price that

[1] For a similar reason, we exclude the 10KG pack size, which ranks 6th and sells sands with pine and wood.

is more than four times higher than or less than one-quarter of the median price for each SKU (across all store-month observations) is treated as an outlier and labeled as missing, following Hitsch et al. (2021). This process filters out unreasonable prices before imputing any missing values of store-SKU-month price (when there is no sale or there is an outlier). For these missing prices, we use the highest observed price from the previous three months for the same SKU at the same store to impute the missing value (Gordon et al., 2013). Finally, we exclude SKUs for a particular store and week if we have never observed them being sold in that store previously.

Next, we aggregate SKUs to create brand-level price indices by converting all SKU prices to a per-unit basis for comparable units. All product categories (e.g., dry food and cat sands) are measured by price per kilogram (KG) except for hygienic mats, which are measured by price per piece. Following the approach of Hitsch et al. (2021), the brand price for brand j at store s in month t, denoted as BrandPrice_{jst}, is calculated as the weighted average price per unit across all SKUs in that brand in the defined market. The weights are based on each SKU's share of offline sales within the brand at that store over the full 30-month data period.

$$\text{Brand\_Price}_{jst} = \sum_{k \in C_{js}} w_{jsk} \times \text{SKU\_Price}_{jstk}, \tag{B.1}$$

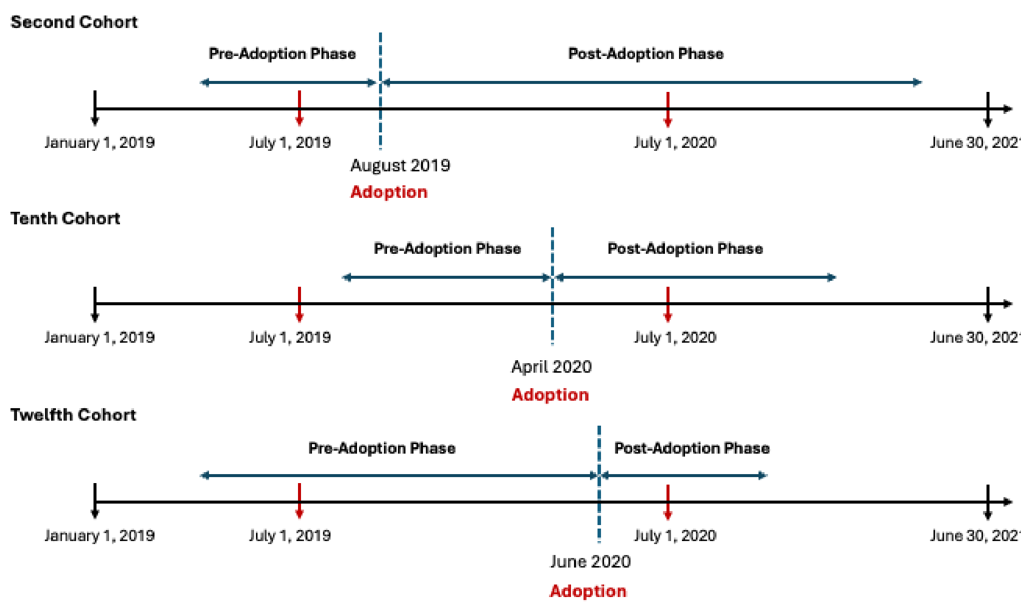
where the weight of SKU k is $w_{jsk} = \frac{\text{Store\_Sales}_{jsk}}{\sum_{k' \in C_{js}} \text{Store\_Sales}_{jsk'}}$, and Store_Sales_{jsk} is the total offline sales of SKU k for brand j at store s from January 1, 2019 to June 30, 2021. C_{js} represents the set of SKUs for brand j available at store s, and SKU_Price_{jstk} is the price of SKU k for brand j sold in store s during month t, as recorded in the store-SKU-month price matrix.

We then apply a similar approach to construct a store-SKU-month level cost matrix based on wholesale cost data, and aggregate SKU-level costs to create brand-level cost indices that correspond to each price index. These cost indices serve as the instrument for addressing price endogeneity, as discussed in Section §3.6. Finally, we use Brazil's Extended National Consumer Index (IPCA) to obtain inflation-adjusted prices and costs (Brazil Institute of Geography and Statistics (IBGE), 2024), with January 2019, the first month of observation, as the base period.

B.1.1.3 Supplementary material for Selection of Customers

Figure B.1 illustrates three examples of adopters defined in §3.3.2, including adopters whose first online purchase in any category during August 2019, April 2020, and June 2020, respectively. Table B.1 presents the sample sizes for each adoption cohort and each category.

Figure B.1. Staggered Adoption Cohorts of Online Shopping (12 Adoption Cohorts from July 2019 to June 2020)



B.1.2 Summary Statistics for Customer Demographics

We summarize the customer demographics for each product category. Table B.2 presents the gender distribution for the different product categories amongst adopters and non-adopters. In all categories, adopters exhibit a higher proportion of women compared to non-adopters. For example, in the dry dog food category, 48% of adopters are female, whereas only 32.6% of non-adopters are female. Figure B.2 shows the cumulative age and income distributions for adopters and non adopters. The height of the leftmost point on each curve represents

Table B.1. Adopters Sample Size by Adoption Time - By Product Category

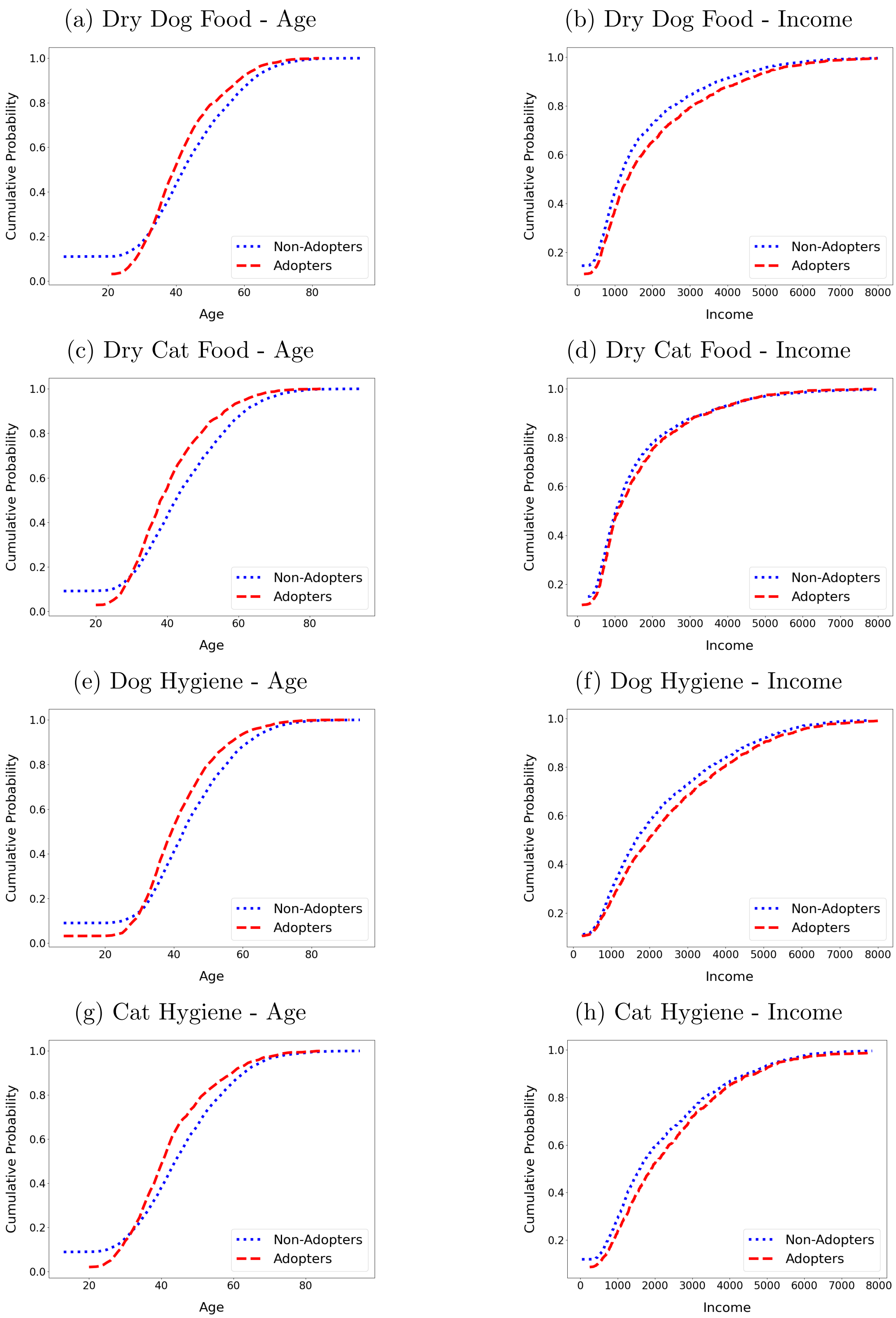
Adoption Month	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
2019-07	85	53	52	36
2019-08	111	46	71	41
2019-09	100	34	96	44
2019-10	110	42	82	47
2019-11	174	51	143	67
2019-12	144	78	120	58
2020-01	138	57	107	47
2020-02	143	52	106	55
2020-03	334	160	382	190
2020-04	438	208	442	241
2020-05	357	143	353	209
2020-06	276	113	260	157
Total	2410	1037	2214	1192

the fraction of customers with missing age or income information. Across all categories, adopters tend to be younger compared to non-adopters. In terms of monthly household income, adopters generally exhibit higher incomes relative to compared to non-adopters.

Table B.2. Gender by Product Categories

	Dry Dog Food		Dry Cat Food		Dog Hygiene		Cat Hygiene	
Gender	Adopters	Non-Adopters	Adopters	Non-Adopters	Adopters	Non-Adopters	Adopters	Non-Adopters
Female	0.480	0.326	0.570	0.418	0.601	0.470	0.602	0.462
Male	0.398	0.522	0.286	0.417	0.285	0.379	0.279	0.387
NaN	0.122	0.152	0.144	0.165	0.113	0.152	0.118	0.151
Total	1	1	1	1	1	1	1	1

Figure B.2. Customer demographics of adopters and non-adopters of product category - age and monthly household income



B.2 Appendix to Descriptive Analysis

B.2.1 Supplementary Result: Percentile Price Paid

As a robustness check for the results in §3.4, Table B.3 presents the within-customer changes in the percentile of prices paid by adopters relative to non-adopters, controlling for customer and year-month fixed effects. The percentile of price paid is calculated by first constructing the empirical price distribution for each month, based on all transactions in that month involving SKUs from the focal category. Each transaction price is then ranked relative to this monthly price distribution to obtain its percentile.

Table B.3. Two-Way Fixed Effects Model Estimates - Percentile_Price_Paid

Dependent Variable:	Percentile_Price_Paid			
	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)	(4)
Variables				
$\mathcal{I}\{\text{Adopter}_i\} \times \mathcal{I}\{t \geq T_i^{\text{adopt}}\}$	-1.437**	-3.318**	-0.9569	-4.104***
	(0.4141)	(0.9727)	(0.6351)	(1.100)
Fixed-effects				
Customer	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes
Fit statistics				
Observations	99,233	27,623	40,183	42,596
R^2	0.86731	0.61190	0.50654	0.42914
Within R^2	0.00057	0.00135	0.00010	0.00126

Clustered (Customer & YearMonth) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

B.3 Appendix to Model and Estimation

B.3.1 Proof for Minorization-Maximization Estimation

Proof of Theorem. We show that $S(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$ is a minorization of $\mathcal{L}(\boldsymbol{\theta}^{(k)})$ at $\boldsymbol{\theta}^{(k)}$ for maximization by checking the following requirements.

- $\mathcal{S}(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)}) \leq \mathcal{L}(\boldsymbol{\theta}^{(k)})$. For a given consumer i at purchase occasion t, we denote the choice probability of consumer i toward brand j during purchase occasion t as Pr_{ijt}:

$$l(\boldsymbol{\psi}_{it};\boldsymbol{y}_{it}) = \prod_j \left(\frac{\exp(\psi_{ijt})}{1+\sum_{j=1}^{J}\exp(U_{ijt})} \right)^{y_{ijt}} \tag{B.2}$$

$$h_j(\boldsymbol{\psi}_{it};\boldsymbol{y}_{it}) = \frac{\partial \log l(\boldsymbol{\psi}_{it};\boldsymbol{y}_{it})}{\partial \psi_{ijt}} = y_{ijt} - Pr_{ijt} \tag{B.3}$$

$$h_{jk}(\boldsymbol{\psi}_{it};\boldsymbol{y}_{it}) = \frac{\partial^2 \log l(\boldsymbol{\psi}_{it};\boldsymbol{y}_{it})}{\partial \psi_{ijt}\partial \psi_{ikt}} = \begin{cases} -Pr_{ijt}(1-Pr_{ikt}), & j=k, \\ Pr_{ijt}Pr_{ikt}, & j \neq k. \end{cases} \tag{B.4}$$

The Taylor expansion of $\log l(\boldsymbol{\psi}_{it})$ at $\tilde{\boldsymbol{\psi}}_{it}$ is

$$\begin{aligned} \log l(\boldsymbol{\psi}_{it};\boldsymbol{y}_{it}) &= \log l(\tilde{\boldsymbol{\psi}}_{it};\boldsymbol{y}_{it}) + (\boldsymbol{\psi}_{it}-\tilde{\boldsymbol{\psi}}_{it})'\nabla \log l(\tilde{\boldsymbol{\psi}}_{it};\boldsymbol{y}_{it}) \\ &\quad + \frac{1}{2}(\boldsymbol{\psi}_{it}-\tilde{\boldsymbol{\psi}}_{it})'\nabla^2 \log l(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it})(\boldsymbol{\psi}_{it}-\tilde{\boldsymbol{\psi}}_{it}). \end{aligned}$$

Where

$$\begin{aligned} \nabla^2 \log l(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) &= \begin{bmatrix} h_{11}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) & h_{12}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) & \cdots & h_{1J}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) \\ h_{21}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) & h_{22}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) & \cdots & h_{2J}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) \\ \vdots & \vdots & \ddots & \vdots \\ h_{J1}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) & h_{J2}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) & \cdots & h_{JJ}(\boldsymbol{\psi}^*_{it};\boldsymbol{y}_{it}) \end{bmatrix} \\ &= \begin{bmatrix} -Pr_{i1t}(1-Pr_{i1t}) & Pr_{i1t}Pr_{i2t} & \cdots & Pr_{i1t}Pr_{iJt} \\ Pr_{i2t}Pr_{i1t} & -Pr_{i2t}(1-Pr_{i2t}) & \cdots & Pr_{i2t}Pr_{iJt} \\ \vdots & \vdots & \ddots & \vdots \\ Pr_{iJt}Pr_{i1t} & Pr_{iJt}Pr_{i2t} & \cdots & -Pr_{iJt}(1-Pr_{iJt}) \end{bmatrix} \end{aligned}$$

We have $\nabla^2 \log l(\boldsymbol{\psi}^*_{it}; \boldsymbol{y}_{it}) \geq -I$, i.e., $\nabla^2 \log l(\boldsymbol{\psi}^*_{it}; \boldsymbol{y}_{it}) + I$ is semi-positive definite.

Therefore,

$$\begin{aligned}\log l(\boldsymbol{\psi}_{it}; \boldsymbol{y}_{it}) &\geq \log l(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it}) + (\boldsymbol{\psi}_{it} - \tilde{\boldsymbol{\psi}}_{it})' \nabla \log l(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it}) - \frac{1}{2}(\boldsymbol{\psi}_{it} - \tilde{\boldsymbol{\psi}}_{it})'(\boldsymbol{\psi}_{it} - \tilde{\boldsymbol{\psi}}_{it}) \\ &= \log l(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it}) - (\tilde{\boldsymbol{\psi}}_{it} - \boldsymbol{\psi}_{it})' \nabla \log l(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it}) - \frac{1}{2}(\tilde{\boldsymbol{\psi}}_{it} - \boldsymbol{\psi}_{it})'(\tilde{\boldsymbol{\psi}}_{it} - \boldsymbol{\psi}_{it}) \\ &= \log l(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it}) - \sum_j h_j(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it})(\tilde{\psi}_{ijt} - \psi_{ijt}) - \frac{1}{2}\sum_j (\tilde{\psi}_{ijt} - \psi_{ijt})^2 \\ &= \log l(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it}) + \frac{1}{2}\sum_j h_j(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it})^2 - \frac{1}{2}\sum_j (\tilde{\psi}_{ijt} + h_j(\tilde{\boldsymbol{\psi}}_{it}; \boldsymbol{y}_{it}) - \psi_{ijt})^2.\end{aligned} \tag{B.5}$$

For now, we derive a lower bound for $\log l(\boldsymbol{\psi}_{it}; \boldsymbol{y}_{it})$, the log-likelihood for each choice instance (i.e., customer-month purchase occasion). The log-likelihood function is its sum over user i and time t. We substitute $\boldsymbol{\psi}_{it}$ and $\tilde{\boldsymbol{\psi}}_{it}$ with the nominal utility $\boldsymbol{x}'_{ijt}\boldsymbol{\alpha} + \delta_{ij}$ and nominal utility evaluated at the current estimates in the k_{th} iteration $\boldsymbol{x}'_{ijt}\boldsymbol{\alpha}^{(k)} + \delta^{(k)}_{ij}$, respectively. Summing over all customers over all periods, the log-likelihood $\mathcal{L}(\boldsymbol{\theta})$ satisfies

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}) &\geq \mathcal{L}(\boldsymbol{\theta}^{(k)}) + \frac{1}{2}\sum_{i,j,t} h_j(\tilde{\boldsymbol{\psi}}^{(k)}_{it}; \boldsymbol{y}_{it})^2 - \frac{1}{2}\sum_{i,j,t}(\tilde{\psi}^{(k)}_{ijt} + h_j(\tilde{\boldsymbol{\psi}}^{(k)}_{it}; \boldsymbol{y}_{it}) - \boldsymbol{x}'_{ijt}\boldsymbol{\alpha} - \delta_{ij})^2 \\ &= \mathcal{L}(\boldsymbol{\theta}^{(k)}) + \frac{1}{2}\sum_{i,j,t} h_j(\boldsymbol{x}'_{ijt}\boldsymbol{\alpha}^{(k)} + \delta^{(k)}_{ij}; \boldsymbol{y}_{it})^2 \\ &\quad - \frac{1}{2}\sum_{i,j,t}(\boldsymbol{x}'_{ijt}\boldsymbol{\alpha}^{(k)} + \delta^{(k)}_{ij} + h_j(\boldsymbol{x}'_{ijt}\boldsymbol{\alpha}^{(k)} + \delta^{(k)}_{ij}; \boldsymbol{y}_{it}) - \boldsymbol{x}'_{ijt}\boldsymbol{\alpha} - \delta_{ij})^2\end{aligned} \tag{B.6}$$

where the RHS of the above inequality is $S(\boldsymbol{\theta}; \boldsymbol{\theta}^{(k)})$.

- $S(\boldsymbol{\theta}^{(k)}; \boldsymbol{\theta}^{(k)}) = \mathcal{L}(\boldsymbol{\theta}^{(k)})$.

By definition,

$$
\begin{aligned}
S(\boldsymbol{\theta}^{(k)};\boldsymbol{\theta}^{(k)}) &= \mathcal{L}(\boldsymbol{\theta}^{(k)}) + \frac{1}{2}\sum_{i,j,t} h_j(\tilde{\boldsymbol{\psi}}_{it}^{(k)};\boldsymbol{y}_{it})^2 - \frac{1}{2}\sum_{i,j,t}(\tilde{\psi}_{ijt}^{(k)} + h_j(\tilde{\boldsymbol{\psi}}_{it}^{(k)};\boldsymbol{y}_{it}) - \boldsymbol{x}_{ijt}'\boldsymbol{\alpha}^{(k)} - \delta_{ij}^{(k)})^2 \\
&= \mathcal{L}(\boldsymbol{\theta}^{(k)}) + \frac{1}{2}\sum_{i,j,t} h_j(\tilde{\boldsymbol{\psi}}_{it}^{(k)};\boldsymbol{y}_{it})^2 - \frac{1}{2}\sum_{i,j,t} h_j(\tilde{\boldsymbol{\psi}}_{it}^{(k)};\boldsymbol{y}_{it})^2 \\
&= \mathcal{L}(\boldsymbol{\theta}^{(k)})
\end{aligned}
\tag{B.7}
$$

B.4 Appendix to Difference-in-Differences (DID) Analysis

B.4.1 Bootstrapped distribution of ATT estimates

Since the customer-month level price elasticities are estimates and not directly observed from data, to obtain correct standard errors, we perform a customer-level block bootstrap with 500 replications. This includes (1) resampling customers with replacement; (2) re-estimating the demand model on each bootstrapped dataset; (3) using the estimated coefficients from that replication to compute price elasticities at the customer-month level (e.g., insignificant demand model coefficients evaluated based on the analytical standard error in the current bootstrap iteration is set to 0); (4) estimating the ATT via staggered DID based on these elasticity estimates; (5) plotting and calculating the bootstrapped standard errors for the ATT and demand parameters.

Figure B.3 presents the bootstrapped distribution of ATT estimates — that is, the effect of online adoption on price elasticity, based on 500 replications.

B.4.2 Two-Way Fixed Effects (TWFE) Estimates

Table B.4 provides TWFE estimates of online adoption on offline price elasticities estimated from the main demand model described in Equation (3.7).

Figure B.3. Bootstrapped distribution of Overall ATT Estimates based on 500 replications with the point estimate Table 3.8.

(a) Dry Dog Food (b) Dog Hygiene (c) Cat Hygiene

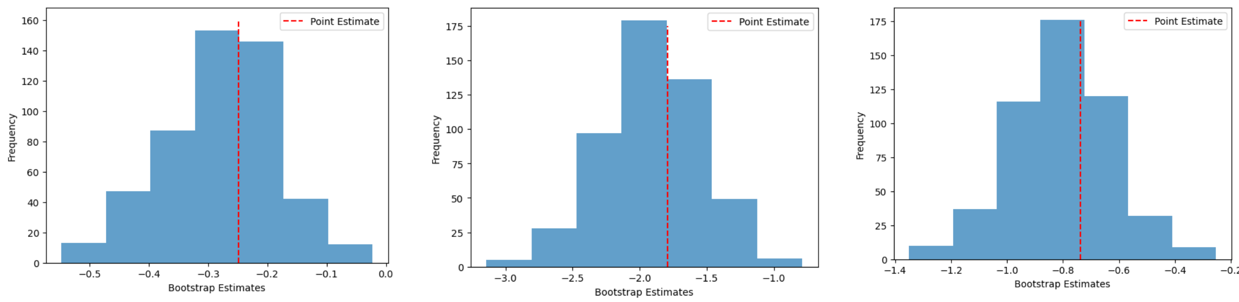


Table B.4. Two-Way Fixed Effects Estimate - Offline Price Elasticity

Dependent Variable:	Offline Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
Variables			
PostOnlineAdoption	-0.2145**	-1.544***	-0.6368***
	(0.0697)	(0.3308)	(0.1605)
Fixed-effects			
CUSTOMER_ID	Yes	Yes	Yes
MONTHSTART	Yes	Yes	Yes
Fit statistics			
Observations	136,411	78,253	55,410
R^2	0.91126	0.76395	0.82317
Within R^2	0.24557	0.32091	0.32582

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

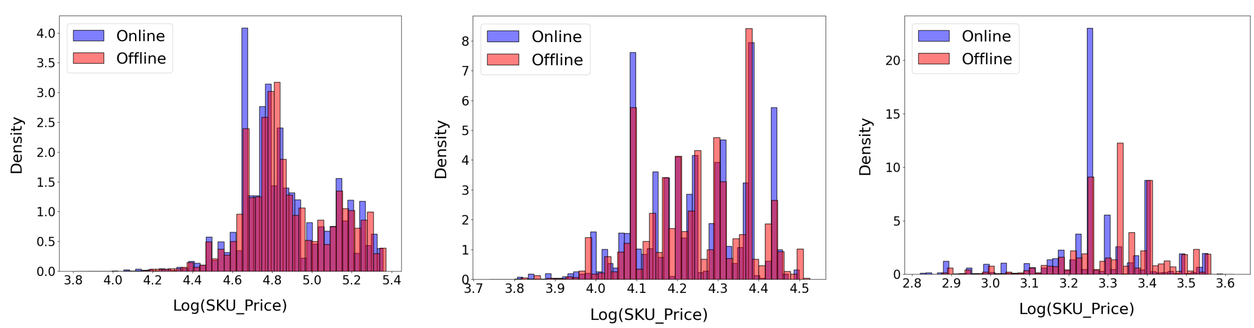
B.5 Appendix to Generalizability and Robustness Checks

B.5.1 Detailed Results for the Role of Price Difference Across Channels

Figure B.4 presents the probability density function (PDF) of offline and online prices (in natural log scale) for each category at the SKU-Store-Date level. While there is substantial price variation across SKUs, the central tendency, skewness, and dispersion of distributions for online and offline channels are quite similar. The substantial overlap in online and offline price distribution suggests no systematic difference between online and offline prices overall.

Figure B.4. Distribution of Online and Offline Prices (Natural Log Scale) over SKU-Date - PDF

(a) Dry Dog Food (b) Dog Hygiene (c) Cat Hygiene



To formally compare offline and online prices, we construct a sample at the SKU-Channel-Store-Date level. Each row represents the price of a specific SKU on a particular channel (online or offline) at a specific store and date. We then estimate the following model:

$$\ln(\text{SKU\_Price}_{kstc}) = \beta \times \text{isOnline}_c + \text{SKU}_k + \text{Date}_t + \epsilon_{kstc} \tag{B.8}$$

where the SKU_Price_{kstc} refers to the selling price of SKU k of store s at time t. The index c refers to the online or offline channel. Here, a negative β would indicate that, on average, online prices are lower than offline prices for the same SKU on the same date, and vice versa. Table B.5 presents the regression result for comparing online and offline prices in Equation (B.8) for three categories. Table B.6 presents the regression results of decomposing SKU price variation for each category. Column (1) of Table B.6 only includes the channel indicator, *isOnline*, and Column (2) includes both channel indicator and SKU fixed effects. Comparing the R-squared values in Table B.6, we see that the price variation across channels is far smaller than the price variation across SKUs.

Table B.5. Descriptive Analysis - Online vs. Offline Prices at the SKU-Date Level

Dependent Variable:	ln(SKU_Price)		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
Variables			
isOnline	-0.0322***	-0.0361*	-0.0607***
	(0.0022)	(0.0135)	(0.0087)
Fixed-effects			
SKU	Yes	Yes	Yes
DATE	Yes	Yes	Yes
Fit statistics			
Observations	554,191	328,323	244,723
R^2	0.95202	0.98493	0.91763
Within R^2	0.01819	0.00912	0.02441

Clustered (SKU & DATE) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

B.5.2 Detailed Results for Alternative Demand Model Specifications

B.5.2.1 Results for Alternate Demand Specification with Binary Adoption Variable

The alternative demand model specification is given by:

$$\begin{aligned}\psi_{ijt} &= \alpha_0 \text{Price}_{jt} + \alpha_1 \text{Post\_OnlineAdoption}_{it} + \alpha_2 \text{Price}_{jt} \times \text{Post\_OnlineAdoption}_{it} \\ &\quad + \alpha_3 \mathcal{I}\{\text{BuyPrevOcca}_{ijt}\} + \beta_1 \text{CF\_Price}_{jt} + \beta_2 \text{CF\_Price\_Post\_OnlineAdoption}_{ijt} + \delta_{ij} \\ &\quad \forall j \in \{1, \ldots, J\}\end{aligned} \tag{B.9}$$

To address the endogeneity of Price_{jt} and its interaction with $\text{Post\_OnlineAdoption}_{it}$, we use the control function approach with instruments Cost_{jt} and $\text{Cost}_{jt} \times \text{Post\_OnlineAdoption}_{it}$,

Table B.6. Descriptive Analysis - SKU-Level Price Variation Decomposition. Compared to column 1, column 2 controls for SKU fixed effects, indicating that after holding SKU constant, the price variation attributable to the online-offline channel is very small.

Dependent Variable:	ln(SKU_Price)					
	Dry Dog Food		Dog Hygiene		Cat Hygiene	
Model:	(1)	(2)	(1)	(2)	(1)	(2)
Variables						
Constant	4.941***		4.043***		3.224***	
	(0.0439)		(0.1420)		(0.0769)	
isOnline	0.0659*	-0.0313***	-0.0123	-0.0382*	-0.0726*	-0.0594***
	(0.0287)	(0.0025)	(0.0269)	(0.0135)	(0.0304)	(0.0087)
Fixed-effects						
SKU		Yes		Yes		Yes
Fit statistics						
Observations	554,191	554,191	328,323	328,323	244,723	244,723
R^2	0.00398	0.89516	1.61×10^{-5}	0.95365	0.00309	0.84228
Within R^2		0.00797		0.00333		0.01235

Clustered (SKU & DATE) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

and construct two control function terms, CF_Price$_{jt}$ and CF_Price_Post_OnlineAdoption$_{ijt}$. We then estimate the above demand model and compute the average elasticity at the customer-month level for each product category following the same steps as before. The results are shown below.

Table B.7 presents the estimates of the demand model. The standard errors for the demand parameters are based on the customer-level block bootstrap with 500 replications.

Table B.7. Coefficient Estimates of Demand Model in Equation (B.9) - Binary Online Adoption Variable

Variables	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Price	-0.3031***	-0.1577***	-1.0463***	-0.2038***
	(0.0171)	(0.0202)	(0.1123)	(0.0199)
Post_OnlineAdoption	-0.0834	-0.2483	2.1377***	1.8521***
	(0.1341)	(0.398)	(0.3641)	(0.3083)
Price * Post_OnlineAdoption	-0.034**	-0.004	-1.0684***	-0.1432***
	(0.0123)	(0.0337)	(0.159)	(0.0218)
BuyPrevOcca	0.1561***	-0.4977***	-0.1802***	0.3216***
	(0.0273)	(0.0451)	(0.0283)	(0.0295)
CF_Price	0.042	0.039	0.0572	0.2797***
	(0.0237)	(0.0268)	(0.1622)	(0.0261)
CF_Price_Post_OnlineAdoption	0.0521**	0.038	0.6873**	0.0609*
	(0.0175)	(0.0343)	(0.2196)	(0.0256)
Customer_Brand_FE	Yes	Yes	Yes	Yes
Observations	136,411	45,443	78,253	55,410
Log-likelihood	-84085.187	-30700.989	-55426.656	-37292.32

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.8 presents the ATT estimates of offline price elasticity from CS Estimator based on the demand model in Equation (B.9). The standard errors for the ATT estimates and the demand model parameters are based on the customer-level block bootstrap with 500 replications; See Figure B.5.

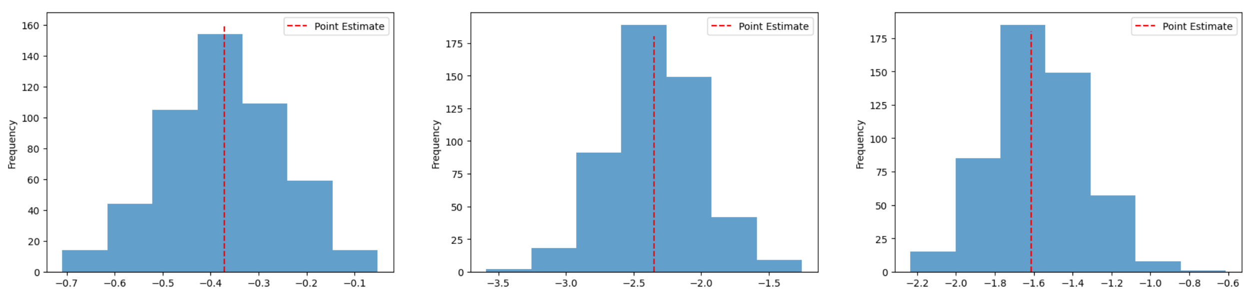
Table B.8. ATT of Online Adoption on Offline Price Elasticity (CS Estimator) - Binary Adoption Variable

Dependent Variable:	Offline Price Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
CS	-0.3711**	-2.3513***	-1.614***
	(0.1224)	(0.3407)	(0.2353)
Observations	136,411	78,253	55,410

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Figure B.5. Bootstrapped distribution of Overall ATT Estimates based on 500 replications with the point estimate from Table B.8.

(a) Dry Dog Food (b) Dog Hygiene (c) Cat Hygiene



B.5.2.2 Results for Alternate Demand Model Controlling for State Dependence due to Online Adoption

The alternative utility specification is given by:

$$\begin{aligned}\psi_{ijt} &= \alpha_0 \text{Price}_{jt} + \alpha_1 \text{cumPercOnlineSpend}_{it} + \alpha_2 \text{Price}_{jt} \times \text{cumPercOnlineSpend}_{it} \\ &\quad + \alpha_3 \mathcal{I}\{\text{BuyPrevOcca}_{ijt}\} + \alpha_4 \mathcal{I}\{\text{BuyPrevOcca}_{ijt}\} \times \text{cumPercOnlineSpend}_{it} \\ &\quad + \beta_1 \text{CF\_Price}_{jt} + \beta_2 \text{CF\_Price\_cumPercOnlineSpend}_{ijt} + \delta_{ij} \quad \forall j \in \{1, \ldots, J\}\end{aligned} \tag{B.10}$$

We estimate this model and again obtain the customer-month level average elasticity for each category as before. Table B.9 presents the estimates of the demand model. The standard errors for the demand model parameters are based on the customer-level block bootstrap with 500 replications.

Table B.9. Coefficient Estimates of Demand Model in Equation (B.10) - Controlling for Interaction between Online Spend and Brand Loyalty

Variables	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Price	-0.3157***	-0.1618***	-1.2497***	-0.2209***
	(0.0159)	(0.0172)	(0.1142)	(0.0186)
cumPercOnlineSpend	-0.3383	-1.5509	8.6901***	3.4632***
	(0.4039)	(1.2012)	(1.8059)	(0.9725)
Price * cumPercOnlineSpend	-0.1092**	0.0227	-4.1435***	-0.284***
	(0.0366)	(0.0994)	(0.7895)	(0.0661)
BuyPrevOcca	0.1592***	-0.5249***	-0.1894***	0.3242***
	(0.0278)	(0.047)	(0.0294)	(0.0306)
CF_Price	0.0521*	0.048	0.1936	0.2803***
	(0.0219)	(0.0245)	(0.1673)	(0.0247)
CF_Price_cumPercOnlineSpend	0.0212	-0.0595	3.6167***	0.3009**
	(0.0719)	(0.1169)	(0.9742)	(0.0915)
BuyPrevOcca * cumPercOnlineSpend	-0.3059	1.0106**	0.2991	0.0133
	(0.2918)	(0.3853)	(0.3022)	(0.2768)
Customer_Brand_FE	Yes	Yes	Yes	Yes
Observations	136,411	45,443	78,253	55,410
Log-likelihood	-84209.589	-30693.216	-55438.072	-37327.394

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.10 presents the ATT estimates of offline price elasticity from the CS estimator based on the demand model in Equation (B.10). The standard errors for the ATT estimates are based on the customer-level block bootstrap with 500 replications; See Figure B.6.

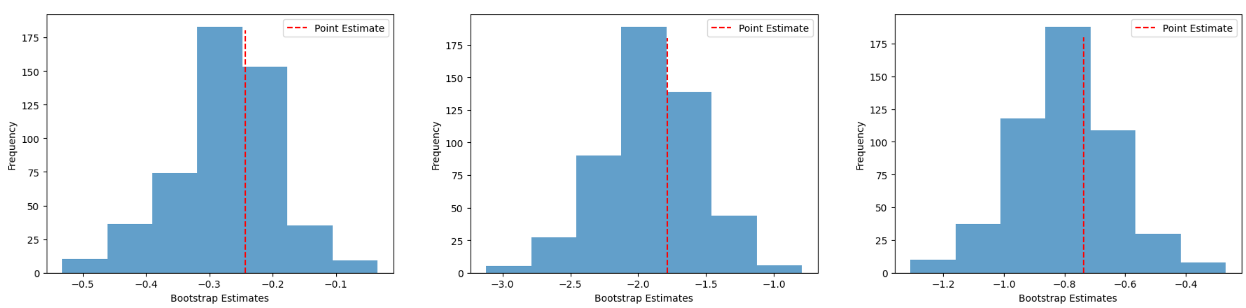
Table B.10. ATT of Online Adoption on Offline Price Elasticity (CS Estimator) - Controlling for State Dependence due to Online Adoption

Dependent Variable:	Offline Price Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
CS	-0.2429**	-1.7868***	-0.7382***
	(0.0797)	(0.3544)	(0.1668)
Observations	136,411	78,253	55,410

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Figure B.6. Bootstrapped distribution of Overall ATT Estimates based on 500 replications with the point estimate from Table B.10.

(a) Dry Dog Food (b) Dog Hygiene (c) Cat Hygiene



B.5.3 Detailed results for testing the Parallel Pre-Trend Assumption

A fundamental condition for the credibility of the estimates presented in §3.7.3 is the parallel trends assumption. We assess it with an event-study figure following Callaway and Sant'Anna (2021). Figure B.7 plots the point estimates and 95% bootstrapped confidence interval for each pre- and post-treatment period, and shows that all pre-treatment coefficients are insignificant from zero. We do not exhibit discernible pre-trends. In addition, the p-values from the Wald tests of the pre-trend assumption are greater than 5% level of significance in each category. We therefore conclude that the parallel trend assumption is reasonably satisfied throughout the analysis.

B.5.4 Detailed Results for Matching, IPW Approaches to Control for Selection on Observables

For each of the adoption cohorts under a product category, we estimate the following propensity score model:

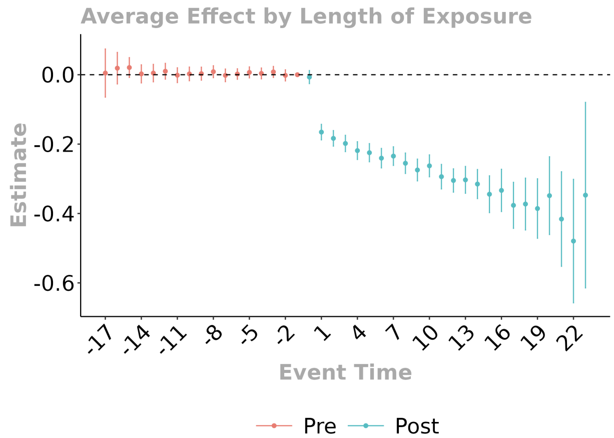
$$\text{PropensityScore}_i = f(X_i^T\theta), \tag{B.11}$$

where $f(\cdot)$ is the logit link function where $f(X_i^T\theta) = \frac{1}{1+\exp(-X_i^T\theta)}$. $\theta \in \mathcal{R}^{M\times 1}$ is a column vector of coefficients. $X_i \in \mathcal{R}^{M\times 1}$ is a column vector with M covariates of customer i, including demographics (e.g., gender, age and income) and the pre-adoption monthly purchase average measured using data up to the adoption period of the adoption cohort. For example, for customers who adopted in July 2019, we measure the pre-adoption variables for both adopters and non-adopters using the data up to June 30, 2019, and estimate the propensity score model for the adoption cohort in July 2019.

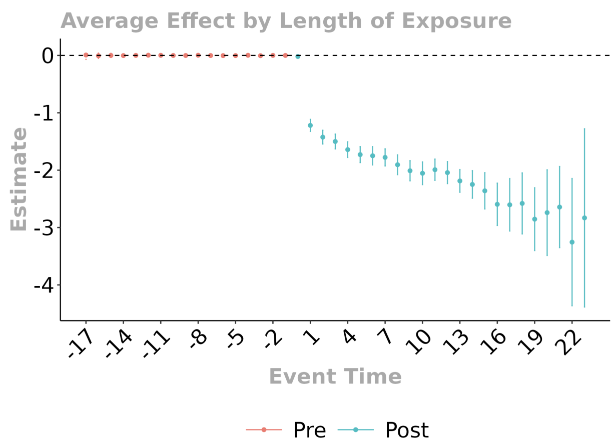
Tables B.11 and B.12 present the propensity score model estimates for the twelve adoption cohorts in the dry dog food, dog hygiene and cat hygiene categories, respectively. Similarly, Tables B.13 and B.14 provide results for dog hygiene and Tables B.15 and B.16 provide results for cat hygiene. We find that certain observables, such as being female, having a high household income, and having a high pre-adoption monthly spend, increase the likelihood of

Figure B.7. Parallel Pre-Trends Evaluation and Dynamic Effect during the Pre- and Post-Adoption Periods

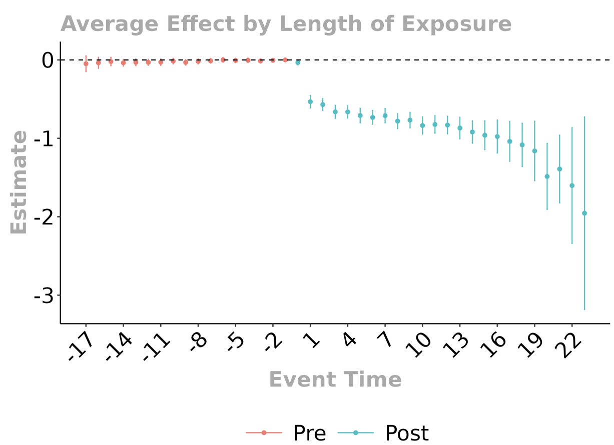
(a) Dry Dog Food



(b) Dog Hygiene



(c) Cat Hygiene



adopting online shopping.

Using the predicted propensity score, we apply the R package *MatchIt* to conduct one-to-one nearest neighbor matching with replacement for each adoption cohort Stuart et al. (2011). For each product category, we first estimate the ATT using the CS based on the matched sample, with results shown in Appendix §B.5.4 Table B.17. Second, we estimate the ATT using the CS estimator with the inverse propensity score weighting (IPW) method, with results shown in Table B.18. These findings are consistent with our main results, demonstrating that consumers become more price sensitive offline after adopting online shopping. Detailed results are below.

B.5.5 ATT Estimates - Using Not-Yet-Treated Adopters as Control Groups

Table B.19 presents ATT estimates using Not-Yet-Treated adopters as the control group, as discussed in §3.8.2.3.

B.5.6 Detailed Results for Alternative Definition of Purchase Occasions

Table B.20 presents the estimates of the main demand model based on an alternative definition of the purchase occasion (i.e., customer-month level during which we observe any interaction with the retailer, either offline or online interaction). The standard errors for the demand parameters are based on the customer-level block bootstrap with 500 replications.

Table B.21 presents the ATT estimates of offline price elasticity from CS Estimator under this alternative definition of purchase occasions and customer-month panel. The standard errors for the ATT estimates and the demand model parameters are based on the customer-level block bootstrap with 500 replications; See Figure B.8.

Table B.11. Propensity Score Logit Model for Dry Dog Food - Adopter Cohorts between 2019-07 and 2019-12 vs. Non-Adopters

	Dependent variable:					
	Adopter					
	2019-07	2019-08	2019-09	2019-10	2019-11	2019-12
Gender_Female	0.288***	0.382***	0.671***	1.030***	0.655***	0.283***
	(0.078)	(0.072)	(0.078)	(0.073)	(0.062)	(0.064)
Gender_NoInformation	−0.944***	−0.177	0.354***	−0.114	0.367***	−0.180
	(0.155)	(0.111)	(0.104)	(0.122)	(0.083)	(0.095)
Age≥50	−0.599***	−0.874***	−0.354*	−0.777***	−1.654***	−1.523***
	(0.166)	(0.175)	(0.160)	(0.153)	(0.111)	(0.112)
Age30_40	0.140	0.667***	0.410**	−0.047	−0.269**	−0.677***
	(0.161)	(0.163)	(0.156)	(0.149)	(0.100)	(0.106)
Age40_50	−0.077	0.134	−0.143	0.204	−1.012***	−0.891***
	(0.163)	(0.165)	(0.161)	(0.145)	(0.106)	(0.107)
Age_NoInformation	−1.798***	−1.188***	−1.496***	−2.105***	−2.036***	−2.120***
	(0.272)	(0.226)	(0.245)	(0.267)	(0.163)	(0.174)
Income<1000	−0.417***	−0.302**	0.038	−0.047	−0.256***	−0.314***
	(0.105)	(0.102)	(0.093)	(0.088)	(0.075)	(0.081)
Income≥2000	0.223*	0.762***	0.149	0.043	0.197**	0.181*
	(0.095)	(0.088)	(0.093)	(0.088)	(0.073)	(0.077)
Income_NoInformation	−0.063	0.060	−0.272*	−0.651***	−0.363***	−0.075
	(0.120)	(0.116)	(0.125)	(0.129)	(0.097)	(0.097)
AvgSpendPerMonth	0.001***	0.001***	0.001***	0.001***	0.0002**	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
AvgQuantitiesPerMonth	−0.095***	−0.108***	−0.058**	0.010*	−0.008	0.010
	(0.025)	(0.024)	(0.018)	(0.005)	(0.008)	(0.009)
AvgUniqueItemsPerMonth	−0.221*	0.231*	−0.507***	−0.444***	−0.235**	−0.166*
	(0.105)	(0.090)	(0.105)	(0.090)	(0.074)	(0.079)
AvgUniqueOrdersPerMonth	0.200**	−0.653***	−0.195*	−0.555***	0.048	−0.027
	(0.075)	(0.102)	(0.084)	(0.102)	(0.062)	(0.071)
AvgUniqueBrandsPerMonth	0.298*	0.148	0.662***	0.778***	0.188	0.365**
	(0.145)	(0.133)	(0.151)	(0.150)	(0.121)	(0.126)
AvgUniqueSubcategoriesPerMonth	0.132	−0.079	1.189***	−0.129	0.243	−0.306
	(0.164)	(0.175)	(0.181)	(0.183)	(0.164)	(0.162)
AvgUniqueCategoriesPerMonth	0.046	0.019	−1.069***	0.172	0.222	0.522***
	(0.142)	(0.159)	(0.160)	(0.158)	(0.144)	(0.155)
Intercept	−2.549***	−2.504***	−2.851***	−2.536***	−1.535***	−1.311***
	(0.178)	(0.180)	(0.177)	(0.165)	(0.117)	(0.118)
Observations	8,023	8,049	8,038	8,048	8,112	8,082
Log Likelihood	−2,586.499	−2,989.092	−2,841.850	−2,963.057	−4,088.072	−3,736.394
Akaike Inf. Crit.	5,206.999	6,012.184	5,717.700	5,960.113	8,210.144	7,506.788
Note:						*p<0.05; **p<0.01; ***p<0.001

Table B.12. Propensity Score Logit Model for Dry Dog Food - Adopter Cohorts between 2020-01 and 2020-06 vs. Non-Adopters

	Dependent variable:					
	Adopter					
	2020-01	2020-02	2020-03	2020-04	2020-05	2020-06
Gender_Female	0.677***	0.473***	0.386***	0.658***	0.736***	0.475***
	(0.065)	(0.064)	(0.048)	(0.044)	(0.047)	(0.050)
Gender_NoInformation	−0.223*	−0.331**	−0.098	0.176**	0.134*	−0.072
	(0.105)	(0.103)	(0.070)	(0.062)	(0.067)	(0.074)
Age≥50	−0.957***	−0.966***	−1.150***	−1.143***	−0.923***	−1.165***
	(0.118)	(0.129)	(0.093)	(0.084)	(0.090)	(0.097)
Age30_40	−0.378**	−0.105	−0.313***	−0.403***	−0.060	−0.113
	(0.115)	(0.122)	(0.090)	(0.083)	(0.088)	(0.092)
Age40_50	−0.682***	−0.363**	−0.698***	−0.602***	−0.602***	−0.742***
	(0.118)	(0.124)	(0.091)	(0.083)	(0.090)	(0.095)
Age_NoInformation	−2.308***	−2.600***	−1.399***	−2.220***	−2.318***	−1.690***
	(0.215)	(0.256)	(0.119)	(0.127)	(0.151)	(0.136)
Income<1000	−0.211*	0.212**	0.051	−0.223***	−0.061	−0.470***
	(0.082)	(0.080)	(0.061)	(0.055)	(0.058)	(0.060)
Income≥2000	0.115	0.163*	0.481***	0.335***	0.210***	−0.334***
	(0.079)	(0.083)	(0.058)	(0.053)	(0.057)	(0.061)
Income_NoInformation	−0.093	−0.149	−0.159*	−0.450***	−0.215**	−0.570***
	(0.100)	(0.106)	(0.078)	(0.071)	(0.074)	(0.078)
AvgSpendPerMonth	0.0004***	0.0002	0.0005***	0.0004***	0.001***	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
AvgQuantitiesPerMonth	−0.051**	−0.008	−0.011	0.009	−0.013	−0.034**
	(0.016)	(0.012)	(0.008)	(0.007)	(0.008)	(0.011)
AvgUniqueItemsPerMonth	−0.285**	−0.557***	−0.156*	−0.076	−0.028	−0.029
	(0.095)	(0.099)	(0.061)	(0.054)	(0.069)	(0.076)
AvgUniqueOrdersPerMonth	−0.223**	−0.409***	−0.558***	−0.455***	−0.786***	−0.267***
	(0.085)	(0.084)	(0.066)	(0.062)	(0.075)	(0.065)
AvgUniqueBrandsPerMonth	0.553***	1.275***	0.286**	−0.134	0.413***	0.547***
	(0.143)	(0.150)	(0.103)	(0.086)	(0.113)	(0.119)
AvgUniqueSubcategoriesPerMonth	0.015	−0.410*	0.467***	0.088	−0.627***	−0.426*
	(0.184)	(0.181)	(0.126)	(0.126)	(0.155)	(0.166)
AvgUniqueCategoriesPerMonth	0.243	0.279	0.025	0.836***	0.907***	0.374*
	(0.165)	(0.178)	(0.121)	(0.122)	(0.139)	(0.149)
Intercept	−1.852***	−2.011***	−1.158***	−0.870***	−1.039***	−0.715***
	(0.132)	(0.136)	(0.099)	(0.091)	(0.099)	(0.101)
Observations	8,076	8,081	8,272	8,376	8,295	8,214
Log Likelihood	−3,577.410	−3,603.534	−6,105.810	−7,120.955	−6,365.540	−5,585.395
Akaike Inf. Crit.	7,188.819	7,241.067	12,245.620	14,275.910	12,765.080	11,204.790
Note:						*p<0.05; **p<0.01; ***p<0.001

Table B.13. Propensity Score Logit Model for Dog Hygiene - Adopter Cohorts between 2019-07 and 2019-12 vs. Non-Adopters

	Dependent variable:					
	Adopter					
	2019-07	2019-08	2019-09	2019-10	2019-11	2019-12
Gender_Female	0.809***	0.697***	0.552***	0.267**	0.436***	0.253**
	(0.118)	(0.105)	(0.089)	(0.092)	(0.075)	(0.080)
Gender_NoInformation	−0.169	0.501***	−0.695***	−0.035	0.076	0.037
	(0.197)	(0.145)	(0.164)	(0.135)	(0.109)	(0.116)
Age≥50	−0.636*	−1.062***	−1.063***	−0.713***	−0.771***	−1.237***
	(0.275)	(0.198)	(0.171)	(0.183)	(0.165)	(0.170)
Age30_40	0.610*	−0.246	−0.244	−0.264	0.077	−0.046
	(0.261)	(0.192)	(0.164)	(0.179)	(0.160)	(0.159)
Age40_50	0.148	−0.400*	−0.705***	−0.442*	−0.103	−0.354*
	(0.264)	(0.194)	(0.167)	(0.179)	(0.160)	(0.161)
Age_NoInformation	−0.059	−1.842***	−2.730***	−1.562***	−0.930***	−2.127***
	(0.303)	(0.302)	(0.358)	(0.266)	(0.208)	(0.277)
Income<1000	−0.354*	−0.558***	−0.183	0.186	0.167	−0.492***
	(0.176)	(0.163)	(0.124)	(0.119)	(0.099)	(0.117)
Income≥2000	0.307*	0.345**	−0.239*	−0.286**	0.074	−0.043
	(0.127)	(0.107)	(0.101)	(0.104)	(0.084)	(0.087)
Income_NoInformation	0.535**	−0.225	0.540***	−0.029	−0.043	0.009
	(0.165)	(0.169)	(0.123)	(0.143)	(0.119)	(0.122)
AvgSpendPerMonth	0.001***	0.001***	0.0004**	0.001***	0.0002	0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
AvgQuantitiesPerMonth	−0.147***	0.015	−0.148***	−0.029	−0.039**	−0.087***
	(0.036)	(0.009)	(0.029)	(0.016)	(0.013)	(0.020)
AvgUniqueItemsPerMonth	−0.043	−0.249***	−0.013	−0.145	−0.327***	−0.149
	(0.115)	(0.072)	(0.089)	(0.078)	(0.071)	(0.092)
AvgUniqueOrdersPerMonth	−0.065	0.187**	−0.684***	−0.455***	−0.149*	−0.103
	(0.091)	(0.069)	(0.095)	(0.091)	(0.061)	(0.092)
AvgUniqueBrandsPerMonth	0.475***	0.462***	0.448***	0.399**	0.425***	−0.253*
	(0.141)	(0.123)	(0.136)	(0.123)	(0.114)	(0.129)
AvgUniqueSubcategoriesPerMonth	0.161	0.023	0.199	0.048	0.655***	1.241***
	(0.167)	(0.161)	(0.169)	(0.151)	(0.141)	(0.178)
AvgUniqueCategoriesPerMonth	−0.463**	−0.239	0.068	0.095	−0.305*	−0.512***
	(0.161)	(0.156)	(0.157)	(0.146)	(0.134)	(0.152)
Intercept	−3.145***	−2.318***	−1.544***	−1.800***	−1.721***	−1.348***
	(0.296)	(0.223)	(0.196)	(0.205)	(0.186)	(0.182)
Observations	3,634	3,653	3,678	3,664	3,725	3,702
Log Likelihood	−1,353.099	−1,680.290	−1,967.035	−1,897.595	−2,712.319	−2,353.311
Akaike Inf. Crit.	2,740.197	3,394.581	3,968.071	3,829.189	5,458.638	4,740.623
Note:						*p<0.05; **p<0.01; ***p<0.001

Table B.14. Propensity Score Logit Model for Dog Hygiene - Adopter Cohorts between 2020-01 and 2020-06 vs. Non-Adopters

	Dependent variable:					
	Adopter					
	2020-01	2020-02	2020-03	2020-04	2020-05	2020-06
Gender_Female	0.627***	0.370***	0.413***	0.582***	0.571***	0.435***
	(0.092)	(0.086)	(0.056)	(0.054)	(0.057)	(0.063)
Gender_NoInformation	0.579***	0.172	−0.093	−0.333***	0.053	−0.056
	(0.123)	(0.121)	(0.083)	(0.085)	(0.084)	(0.093)
Age≥50	−1.372***	−1.130***	−1.080***	−0.834***	−1.256***	−1.169***
	(0.171)	(0.168)	(0.114)	(0.119)	(0.121)	(0.128)
Age30_40	−0.134	−0.341*	−0.483***	−0.122	−0.318**	−0.274*
	(0.157)	(0.162)	(0.114)	(0.118)	(0.118)	(0.123)
Age40_50	−0.802***	−0.800***	−0.964***	−0.261*	−0.675***	−0.825***
	(0.163)	(0.165)	(0.114)	(0.117)	(0.119)	(0.125)
Age_NoInformation	−1.945***	−2.592***	−2.108***	−1.115***	−1.837***	−2.046***
	(0.252)	(0.286)	(0.162)	(0.148)	(0.160)	(0.184)
Income<1000	0.277*	−0.157	−0.118	−0.144	−0.222**	−0.182*
	(0.118)	(0.125)	(0.081)	(0.077)	(0.081)	(0.090)
Income≥2000	0.304**	0.246**	0.430***	0.331***	0.233***	0.263***
	(0.100)	(0.095)	(0.062)	(0.060)	(0.062)	(0.069)
Income_NoInformation	−0.080	0.119	−0.263**	−0.317***	−0.059	−0.137
	(0.144)	(0.134)	(0.096)	(0.092)	(0.093)	(0.105)
AvgSpendPerMonth	0.001***	0.001***	0.001***	0.001***	0.001***	0.0005***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
AvgQuantitiesPerMonth	−0.150***	−0.007	0.011	−0.046***	−0.009	−0.014
	(0.024)	(0.007)	(0.008)	(0.012)	(0.009)	(0.009)
AvgUniqueItemsPerMonth	0.140	−0.345***	−0.139**	0.056	−0.185**	0.138*
	(0.085)	(0.075)	(0.054)	(0.065)	(0.058)	(0.070)
AvgUniqueOrdersPerMonth	−0.124	−0.080	−0.172***	−0.498***	−0.372***	−0.605***
	(0.084)	(0.088)	(0.038)	(0.064)	(0.067)	(0.084)
AvgUniqueBrandsPerMonth	−0.050	0.436***	0.087	−0.224*	0.242**	−0.292*
	(0.133)	(0.113)	(0.091)	(0.102)	(0.093)	(0.116)
AvgUniqueSubcategoriesPerMonth	0.951***	0.664***	0.487***	0.922***	0.502***	0.919***
	(0.171)	(0.156)	(0.116)	(0.126)	(0.131)	(0.147)
AvgUniqueCategoriesPerMonth	−0.616***	−0.769***	−0.139	−0.257*	−0.264*	−0.394**
	(0.154)	(0.160)	(0.113)	(0.117)	(0.135)	(0.139)
Intercept	−2.091***	−1.383***	−0.282*	−0.497***	−0.247	−0.426**
	(0.196)	(0.190)	(0.125)	(0.134)	(0.132)	(0.141)
Observations	3,689	3,688	3,964	4,024	3,935	3,842
Log Likelihood	−2,067.730	−2,175.618	−4,499.026	−4,832.667	−4,400.367	−3,656.464
Akaike Inf. Crit.	4,169.460	4,385.235	9,032.052	9,699.334	8,834.733	7,346.928

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.15. Propensity Score Logit Model for Cat Hygiene - Adopter Cohorts between 2019-07 and 2019-12 vs. Non-Adopters

	Dependent variable:					
	Adopter					
	2019-07	2019-08	2019-09	2019-10	2019-11	2019-12
Gender_Female	0.835***	0.450***	0.986***	0.823***	0.526***	0.364***
	(0.141)	(0.129)	(0.130)	(0.131)	(0.107)	(0.107)
Gender_NoInformation	0.055	0.315	−0.655*	0.164	0.034	−0.424*
	(0.229)	(0.179)	(0.259)	(0.195)	(0.158)	(0.179)
Age≥50	−0.024	−0.217	−0.605*	−0.916***	−0.503	−0.764**
	(0.361)	(0.353)	(0.277)	(0.218)	(0.264)	(0.236)
Age30_40	1.468***	1.077**	0.730**	−0.334	1.052***	0.182
	(0.352)	(0.345)	(0.270)	(0.219)	(0.254)	(0.234)
Age40_50	0.302	0.325	0.057	−0.929***	0.504*	0.031
	(0.361)	(0.349)	(0.269)	(0.222)	(0.255)	(0.229)
Age_NoInformation	−15.840	−0.103	−16.557	−17.598	−1.475***	−1.916***
	(380.738)	(0.404)	(382.142)	(380.148)	(0.408)	(0.394)
Income<1000	0.269	−0.407*	−0.334*	−0.417*	−0.544***	−0.396*
	(0.229)	(0.182)	(0.161)	(0.199)	(0.153)	(0.171)
Income≥2000	1.356***	−0.137	−0.495***	0.368**	0.076	0.440***
	(0.171)	(0.134)	(0.126)	(0.130)	(0.110)	(0.117)
Income_NoInformation	0.072	−0.148	−1.423***	−0.066	−0.364*	−0.507*
	(0.281)	(0.187)	(0.255)	(0.193)	(0.171)	(0.200)
AvgSpendPerMonth	0.001**	0.001***	0.001***	0.002***	−0.0001	0.0002
	(0.0002)	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0002)
AvgQuantitiesPerMonth	0.018*	−0.003	0.005	−0.076***	0.044***	0.009
	(0.008)	(0.008)	(0.009)	(0.015)	(0.006)	(0.008)
AvgUniqueItemsPerMonth	−0.213*	0.101	−0.104	0.241**	−0.649***	−0.268**
	(0.090)	(0.053)	(0.079)	(0.078)	(0.094)	(0.082)
AvgUniqueOrdersPerMonth	−0.076	−0.176*	−0.340**	−0.804***	−0.962***	−0.716***
	(0.109)	(0.084)	(0.104)	(0.116)	(0.126)	(0.115)
AvgUniqueBrandsPerMonth	0.116	−0.003	0.143	−0.528***	0.813***	−0.002
	(0.186)	(0.117)	(0.138)	(0.157)	(0.172)	(0.159)
AvgUniqueSubcategoriesPerMonth	0.417	−0.514**	−0.142	−0.009	0.608**	−0.093
	(0.225)	(0.160)	(0.206)	(0.206)	(0.207)	(0.193)
AvgUniqueCategoriesPerMonth	−0.786***	0.757***	0.394*	0.798***	−0.209	1.401***
	(0.228)	(0.167)	(0.198)	(0.211)	(0.196)	(0.194)
Intercept	−3.263***	−3.235***	−2.481***	−2.102***	−2.167***	−2.533***
	(0.401)	(0.371)	(0.299)	(0.253)	(0.268)	(0.256)
Observations	2,761	2,766	2,769	2,772	2,792	2,783
Log Likelihood	−908.558	−1,079.992	−1,078.742	−1,114.206	−1,440.825	−1,336.839
Akaike Inf. Crit.	1,851.117	2,193.985	2,191.483	2,262.411	2,915.651	2,707.678

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.16. Propensity Score Logit Model for Cat Hygiene - Adopter Cohorts between 2020-01 and 2020-06 vs. Non-Adopters

	Dependent variable:					
	Adopter					
	2020-01	2020-02	2020-03	2020-04	2020-05	2020-06
Gender_Female	0.798***	0.887***	0.490***	0.459***	0.357***	0.702***
	(0.131)	(0.122)	(0.071)	(0.068)	(0.068)	(0.078)
Gender_NoInformation	0.835***	0.334	−0.072	0.224*	−0.370***	0.232*
	(0.164)	(0.176)	(0.108)	(0.094)	(0.106)	(0.111)
Age≥50	−0.962***	−1.164***	−0.881***	−0.896***	−1.128***	−1.857***
	(0.253)	(0.202)	(0.141)	(0.130)	(0.132)	(0.133)
Age30_40	0.505*	−0.194	−0.145	−0.237	−0.319*	−0.801***
	(0.240)	(0.195)	(0.142)	(0.131)	(0.133)	(0.130)
Age40_50	−0.014	−0.816***	−0.727***	−0.629***	−0.890***	−1.147***
	(0.240)	(0.199)	(0.141)	(0.130)	(0.133)	(0.128)
Age_NoInformation	−0.407	−2.189***	−2.352***	−2.549***	−1.864***	−3.821***
	(0.298)	(0.370)	(0.237)	(0.231)	(0.190)	(0.344)
Income<1000	−0.463**	0.008	−0.241*	0.499***	−0.160	−0.097
	(0.174)	(0.140)	(0.104)	(0.091)	(0.095)	(0.107)
Income≥2000	0.122	−0.463***	0.428***	0.465***	0.038	0.307***
	(0.124)	(0.123)	(0.077)	(0.075)	(0.074)	(0.082)
Income_NoInformation	−0.508*	−0.588**	−0.516***	0.224*	−0.367**	−0.214
	(0.205)	(0.194)	(0.131)	(0.108)	(0.113)	(0.128)
AvgSpendPerMonth	−0.002***	0.0004*	−0.00003	0.0003**	−0.00001	−0.0002
	(0.0003)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
AvgQuantitiesPerMonth	0.031**	0.012	0.008	−0.001	0.008	0.012
	(0.012)	(0.010)	(0.006)	(0.006)	(0.006)	(0.007)
AvgUniqueItemsPerMonth	−0.673***	−0.410***	−0.090*	−0.029	−0.072	−0.146*
	(0.108)	(0.091)	(0.040)	(0.038)	(0.045)	(0.058)
AvgUniqueOrdersPerMonth	−0.330**	−0.696***	−0.214***	−0.482***	−0.512***	−0.174**
	(0.119)	(0.112)	(0.060)	(0.068)	(0.066)	(0.058)
AvgUniqueBrandsPerMonth	0.586**	0.323*	0.167	0.006	−0.135	0.205
	(0.192)	(0.157)	(0.085)	(0.085)	(0.093)	(0.111)
AvgUniqueSubcategoriesPerMonth	0.745**	0.984***	0.073	−0.027	0.787***	0.762***
	(0.253)	(0.169)	(0.115)	(0.095)	(0.128)	(0.157)
AvgUniqueCategoriesPerMonth	0.231	−0.441**	0.326**	0.642***	−0.109	−0.721***
	(0.240)	(0.161)	(0.117)	(0.111)	(0.127)	(0.162)
Intercept	−2.797***	−1.602***	−0.789***	−0.740***	−0.040	−0.024
	(0.274)	(0.220)	(0.157)	(0.147)	(0.142)	(0.151)
Observations	2,772	2,780	2,915	2,966	2,934	2,882
Log Likelihood	−1,163.361	−1,274.134	−2,813.474	−3,215.689	−3,002.450	−2,487.669
Akaike Inf. Crit.	2,360.722	2,582.269	5,660.947	6,465.378	6,038.899	5,009.338

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.17. ATT of Online Adoption on Offline Price Elasticity (CS Estimator) - Propensity Score Matching based on each adoption cohort's time of adoption

Dependent Variable:	Offline Price Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
CS	-0.2473**	-1.788***	-0.7409***
	(0.0945)	(0.3678)	(0.1794)
Observations	58,580	50,913	29,039

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

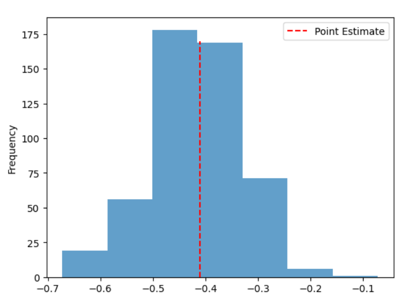
Table B.18. ATT of Online Adoption on Offline Price Elasticity (CS Estimator) - IPW

Dependent Variable:	Offline Price Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
CS	-0.2479**	-1.7903***	-0.7399***
	(0.0945)	(0.3678)	(0.1793)
Observations	136,411	78,253	55,410

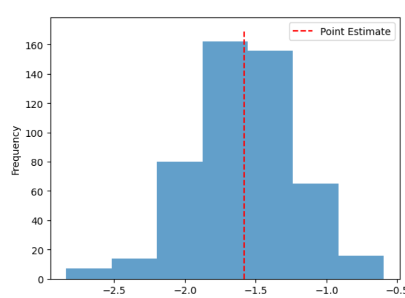
Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Figure B.8. Bootstrapped distribution of Overall ATT Estimates based on 500 replications with the point estimate from Table B.21.

(a) Dry Dog Food



(b) Dog Hygiene



(c) Cat Hygiene

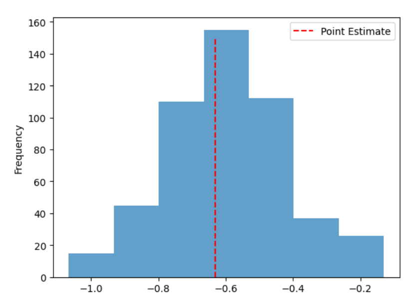


Table B.19. ATT of Online Adoption on Offline Price Elasticity (CS Estimator) - Only Use Not-Yet-Treated Adopters as Comparison Units

Dependent Variable:	Offline Price Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
CS	-0.1685*	-1.1729***	-0.5222***
	(0.0814)	(0.3006)	(0.1533)
Observations	29,416	28,375	15,081

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.20. Coefficients of Demand Model Estimate - Alternative Definition of Purchase Occasions

Variables	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Price	-0.3171***	-0.1637***	-1.4982***	-0.2933***
	(0.0204)	(0.0185)	(0.1166)	(0.0205)
cumPercOnlineSpend	-1.1362**	-2.8399*	4.822**	0.8407
	(0.365)	(1.118)	(1.6143)	(0.9654)
Price * cumPercOnlineSpend	-0.1354***	0.0413	-2.8394***	-0.1615*
	(0.0333)	(0.0926)	(0.7037)	(0.0652)
BuyPrevOcca	0.173***	-0.4511***	-0.1633***	0.3556***
	(0.0263)	(0.0437)	(0.0275)	(0.0295)
CF_Price	0.0421	0.0596*	0.4807**	0.3616***
	(0.0257)	(0.0247)	(0.1714)	(0.0272)
CF_Price_cumPercOnlineSpend	-0.01	-0.1603	2.372**	0.1731
	(0.0623)	(0.1016)	(0.881)	(0.0889)
Customer_Brand_FE	Yes	Yes	Yes	Yes
Observations	144,094	48,737	84,689	57,344
Log-likelihood	-88735.39	-32496.48	-58287.349	-38561.677

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.21. ATT of Online Adoption on Offline Price Elasticity (CS Estimator) – Alternative Definition of Purchase Occasions

Dependent Variable:	Offline Price Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
ATT	-0.4104***	-1.5819***	-0.6315***
	(0.0867)	(0.3688)	(0.1799)
Observations	144,094	84,689	57,344

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

B.5.7 Detailed Results for Alternative Demand Equation with Control Function for Price Term Only

In this section, we follow Ebbes et al. (2016) to implement the control function method by running Equation (3.4) and adding the control function term only for price. We show that our main findings are consistent when only adding the control function terms for price. Table B.22 presents the estimates of the main demand model with control function for price term only. The standard errors for the demand parameters are based on the customer-level block bootstrap with 500 replications.

Table B.23 presents the ATT estimates of offline price elasticity when we estimate the demand model in Equation (3.7) with a control function term for price only. The standard errors for the ATT estimates and the demand model parameters are based on the customer-level block bootstrap with 500 replications; See Figure B.9.

Table B.22. Coefficient Estimates - Demand Model with Single Control Function Term for Price

Variables	Dry Dog Food	Dry Cat Food	Dog Hygiene	Cat Hygiene
Price	-0.3165***	-0.1588***	-1.3137***	-0.2401***
	(0.0212)	(0.018)	(0.1163)	(0.0197)
cumPercOnlineSpend	-0.3837	-1.1528	3.7018**	2.4235*
	(0.4008)	(1.1453)	(1.1297)	(0.9528)
Price * cumPercOnlineSpend	-0.1076**	-0.0017	-1.9529***	-0.2061**
	(0.0374)	(0.0974)	(0.4989)	(0.0638)
BuyPrevOcca	0.151***	-0.5032***	-0.1805***	0.326***
	(0.0272)	(0.045)	(0.0283)	(0.0294)
CF_Price	0.0524	0.0432	0.3133	0.3093***
	(0.0268)	(0.0247)	(0.1697)	(0.0257)
Customer_Brand_FE	Yes	Yes	Yes	Yes
Observations	136,411	45,443	78,253	55,410
Log-likelihood	-84211.483	-30698.359	-55455.669	-37337.88

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.23. ATT of Online Adoption on Offline Price Elasticity (CS Estimator) – Including Single CF Term for Price Endogeneity

Dependent Variable:	Offline Price Elasticity		
	Dry Dog Food	Dog Hygiene	Cat Hygiene
Model:	(1)	(2)	(3)
ATT	-0.2394**	-0.867***	-0.5584***
	(0.0807)	(0.2119)	(0.1566)
Observations	136,411	78,253	55,410

Bootstrapped standard-errors in parentheses via 500 replications
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Figure B.9. Bootstrapped distribution of Overall ATT Estimates based on 500 replications with the point estimate from Table B.23.

(a) Dry Dog Food (b) Dog Hygiene (c) Cat Hygiene