Is Online Retail Killing Coffee Shops?
Estimating the Winners and Losers of Online Retail Using Customer Transaction Microdata

Lindsay E. Relihan*
Purdue University

August 2022
(first version November 2017)

Abstract

Is online retail a complement or substitute to local offline economies? This paper provides the first evidence on mechanisms that create offline winners to online retail. I find the most important is the time savings of online retail, which consumers use to increase their trips for time-intensive services like coffee shops. I use new, detailed data on the daily transactions of millions of anonymized customers. I then estimate a discrete choice model of consumer trip choice which accounts for correlations in trip utility shocks. I show that the model matches key features of observed behavior that are missed by more standard models, such as the disproportionate increase in trips to nearby coffee shops when consumers switch to online groceries. Model counterfactuals are used to forecast changes in future trip demand and outline strategies which offline retailers can use to compete against online retail. For consumers, I find that the welfare gains from online grocery platforms go disproportionately to high-income consumers.

*Contact by email at lrelihan@purdue.edu or by mail at Department of Economics, Purdue University, 340 W. State Street, West Lafayette, IN 47906. I thank Gilles Duranton, Jessie Handbury, Joseph Gyourko, Todd Sinai, Benjamin Keys, Daniel Sturm, Henry Overman, and Vernon Henderson for many helpful comments. I also thank seminar participants at LSE, Wharton, University of Illinois, USC, UC San Diego, University of Utah, Baruch College, the Federal Home Loan Mortgage Corporation, Federal Reserve Board, Federal Reserve Bank of Philadelphia, Rice University, Hebrew University, the UEA Online Spatial and Urban Seminar, and the 2017 and 2021 meetings of the North American UEA. I also thank colleagues at JPMorgan Chase for their comments and Mingchun Hu for excellent research assistance. I also gratefully acknowledge supporting funding received through the Benjamin H. Stevens Graduate Fellowship in Regional Science and the North American Regional Science Council. This research was made possible by a data use agreement between myself and the JPMorgan Chase Institute (JPMCI), which has created de-identified data assets that are selectively available to be used for academic research. More information about JPMCI de-identified data assets and data privacy protocols are available at www.jpmorganchase.com/institute. All statistics from JPMCI data reflect cells with at least 10 observations. The opinions expressed are my own and do not represent the views of JPMorgan Chase & Co. While working on this paper, I was a paid contractor of JPMCI.

1 Introduction

Driven by the COVID-19 pandemic, U.S. households increased their online retail purchases 31.9% between the first and second quarters of 2020 and have continued to spend online at similarly high rates.[1] This has the potential to accelerate the longer-term shift to online retail as consumers gain familiarity with online products and adopt new consumption behaviors. Consequently, the "retail apocalypse" for many traditional brick-and-mortar stores is intensifying and, regardless of the pandemic's trajectory, many are likely to close permanently.[2] However, this narrative ignores that, like other products, there are likely to be both complements *and* substitutes to online retail. Therefore, some brick-and-mortar stores might survive, or even benefit, from the rise of online retail in the long-run. Insight into the mechanisms that create these complements and their welfare consequences is vital to understanding the future of local economic activity and the strategies that best support offline economies.

In this paper, I provide the first empirical evidence that online retail can create both winning and losing brick-and-mortar stores. I explore multiple mechanisms which could create these winners and pinpoint changes in time use as a key driver. One of the major benefits of online retail is the time saved physically traveling to and at stores (Forman et al. 2009). Consumers who shop online, therefore, gain time that they can substitute toward new trips. Furthermore, consumers' trip-chaining behavior can create winners and losers even among store types that benefit overall from the time-savings of online retail.[3] For example, a consumer who visits a coffee shop because it is close to a grocery store may visit a coffee shop closer to home instead when they purchase groceries online. Thus, some stores can be offline shopping complements to online retail if time use preferences cause a reorganization of consumers' shopping trips when they purchase more online. Previous research has been unable to study whether such fine-grained effects exist because detection requires both a large, detailed dataset and an empirical identification strategy for causal effects.

I use data consisting of the daily credit and debit card transactions of tens of millions of anonymized customers from JPMorgan Chase from October 2012 - May 2017. With these data, I provide new summary statistics on trip formation and features of the online grocery market. During the mid-2010s, online grocery platforms entered cities in the US in quick succession, generating quasi-random variation in entry. However, as is the case with many new products, few customers used online groceries initially, such that city-wide effects of platforms were negligible. Therefore, I study the effects of online groceries on shopping trips for a subset of customers for whom exogenous platform entry plausibly induced exogenous platform adoption timing: early platform adopters. For these customers, I compare the changes in their trips before and after platform adoption against

[1] According to the Census Monthly Retail Trade Survey, typical pre-pandemic 2-year-over-year growth rates for online retail hovered around 30% for the last decade. Since the pandemic began, those growth rates have doubled.

[2] Recent work shows that sectors exposed to social distancing were hit hard, such as restaurants, with other sectors, like clothing, also experiencing more closures (Crane et al. 2021).

[3] Trip-chaining is defined as grouping visits to multiple locations on one trip to save on travel costs and the fixed cost of making a trip.

the trip choices of a set of matched platform non-users from their same neighborhoods as controls.

I first show that early adopters of online grocery platforms make broad changes in their store choices beyond reducing their visits to grocery stores. Before adoption, early platform adopters visit grocery stores 6.8 days per month. In the year after adoption, they reduce that frequency by an average of 0.5 days per month, a reduction of 7.9%. At the same time, I find that early adopters increase their spending on local ground transportation by 12.9%, consistent with consumers making new trips when they become online grocery shoppers. I find that they increase their trips including services, such as coffee shops, leisure goods and services, personal care services, and restaurants. But for most goods, including clothing, general goods, and home goods, I find no such increase. This difference suggests that consumers use the time saved from shopping at grocery stores to increase activities that require substantial time to consume, making time use key in creating winners to online retail.

To further support the role of time use, I make a more detailed study of consumers' bundled trip choices for a grocery store and/or coffee shop. I focus on coffee shops for illustration, as they are one service unlikely to be affected by additional mechanisms. I find strong day-of-week differences in effects consistent with different opportunity costs of time. For example, consumers primarily substitute away from trips including the grocery store toward trips to neither store the most on Sundays, when non-shopping activities like visits to friends and family might be more attractive. In contrast, consumers increase their trips only to coffee shops the most during the week versus the weekend, when busy workers might particularly benefit from more coffee.

I also find evidence that distance impacts the extent to which different stores win or lose from online grocery adoption. For example, consumers reduce their chained trips to a grocery store and coffee shop, but increase such trips for stores that are located close together. Moreover, I find that consumers increase their trips only to coffee shops more for coffee shops close to consumers' homes. These two effects show that while grocery stores lose and coffee shops win overall, distance costs affect the extent to which different stores win or lose from online retail competition.

Standard discrete choice models that carry independence of irrelevant alternatives (IIA) assumptions cannot capture many of these substitution patterns. Therefore, I build a discrete choice model of consumer trip choice that relaxes the IIA assumption through the inclusion of separate substitutability parameters for each pair of trips. I then estimate the model using the same panel of early platform adopters and matched non-users, allowing me to leverage the same variation in adoption timing to identify the model's parameters. For instance, trip pair substitutability parameters are identified from trip choices at different distance costs and changes in trip choices following platform adoption. Simulations show that the model outperforms standard models in predicting trip substitution patterns consistent with the reduced form.

I then use the model to estimate the welfare gain to consumers of online groceries and the impact of strategies offline firms can use to compete with a larger online grocery market. I find that the estimated welfare gains for consumers to online grocery platforms are strongly associated with income – welfare gains in the highest zip code median income quintile are three times higher than the lowest quintile. However, in general, welfare gains are small during my sample period due to initially low adoption rates across the population. I show that increases in platform value representing a more mature market have the potential to markedly increase platform adoption rates. For a 50% increase in platform values, mean adoption rates across zip codes jump from 1.7% to 9.3%. As a result, I predict that mean zip code grocery store trip frequency would fall by 1.8%, while the mean frequency for coffee shops would rise by 2.8%.

In counterfactual exercises, I measure the effects of strategies that offline grocery stores could use to compete in a market with higher online grocery platform adoption. Because the online platform adoption population is likely to remain relatively modest and customers who adopt them only partially replace offline trips to grocery stores, strategies which increase the benefits of chained trips, physical access, and store value have the potential to substantially limit or reverse negative effects. However, these strategies are likely to be less successful where retail services and alternative activities are plentiful, such as downtowns, because consumers benefit more from these activities when they replace offline trips with online purchases.

This paper closely relates to research exploring the effects of online retail on consumer consumption and welfare. Most relevant is the work focused on online versus offline shopping behavior and its implications for retail firm strategies, both within and across channels (Gentzkow 2007; Brynjolfsson et al. 2009; Avery et al. 2012; Pozzi 2012). A growing body of work documents additional sources of welfare gains to online retail, including reductions in search costs (Bakos 1997; Jin and Kato 2007), trade frictions (Jingting et al. 2018; Couture et al. 2020) and prices (Jo et al. 2019) and increases in product variety (Quan and Williams 2018). Dolfen et al. (2019) use similar card transaction data to quantify the relative importance of several of these channels. They also find that welfare gains from new online products are higher for high-income consumers and those in urban areas. Also related is work measuring the substitution effects of online retail using changes in online sales taxes. Examples include (Goolsbee 2000; Ellison and Ellison 2009; Einav et al. 2014; Baugh et al. 2018). This paper uses a new natural experiment to study the effect of online retail and mechanisms that impact both welfare and firm strategy.

The results also relate to work studying the consumption value of cities (Glaeser et al. 2001). For firms, we know that location decisions are shaped by spatial competition (Hotelling 1929; Serra and Colomé 2001; Houde 2012; Ushchev et al. 2015) and the benefits of retail agglomeration (Arentze et al. 2005; Brandão et al. 2014; Jardim 2015). Recent research finds that the value of urban density is large and based on access to non-tradable products and services (Glaeser et al. 2001; Handbury and Weinstein 2015; Cosman 2017; Couture 2016; Davis et al. 2019; Gorback 2020; Couture et al. 2021). In this vein, Su (2022) shows that time use is an important dimension to the consumption of these urban amenities. The finding that consumers use additional time from online retail to purchase more services is also in line with decades long trends in allocating free time toward leisure activities (Aguiar and Hurst 2007b). This research suggests that in response to online retail, firms will co-locate more and closer to consumers and supports the view that cities can be a complement to online retail (Sinai and Waldfogel 2004).

More broadly, this research contributes to our understanding of the impact of shopping behavior on consumption choices. To date, differences in consumer's opportunity cost of time generated by life-cycle and business-cycle changes have been used to study substitution between time spent shopping and consumption (Aguiar and Hurst 2007a; Aguiar et al. 2013; Nevo and Wong 2019; Bronnenberg et al. 2020). Long acknowledged is the importance of travel costs in determining consumer store choices (Narula et al. 1983; Harwitz et al. 1983). However, the impact of trip-chaining on those choices has so far mainly been studied in the marketing literature (Dellaert et al. 1998; Brooks et al. 2004, 2008), with the notable exceptions of Baker et al. (2020) and Miyauchi et al. (2022) who also find that trip fixed costs can create complementary purchases.

2 Customer Transaction Data

I use proprietary data from JPMorgan Chase (JPMC) containing 53 billion transactions made by 69 million anonymized customers over October 2012 - May 2017.[4] The transactions data contain important details that are vital to studying variation in consumer spending across channel, time, and space. In addition to basic information such as the date and transaction amount, I use fields attached to each transaction to determine the locations, payment channel, merchant, store, and product. Locations for both customers and merchants are at the zip code level.[5] Payment channel is determined by whether the card was present at the time of purchase, the merchant is a known online-only retailer, or the store location is characterized by a website or phone number. Merchants and stores are identified via regular expression matching against the description for each transaction and the store zip code. Each transaction also carries a four-digit merchant classification code which characterizes the good or service sold. I define a product set with a mapping from these classification codes, complemented by regular expression matching on key merchants where these codes are inadequate.[6]

There are a number of features of the data that set it apart from other available datasets. One is the shear size. As discussed below, this size is key to studying new or small markets, like that for online groceries, and for attaining adequate coverage at fine geographic areas. This size requirement

[4] I focus on customers over households because the latter are difficult to characterize in the data. Customers which share an address are likely part of the same household, but to fully capture households each member must be a customer of JPMC. Some customers in the data may reflect the spending of multiple people – for example, customers with joint accounts.

[5] Customer zip codes are based on customer mailing addresses. Merchant zip codes are based on card terminal locations.

[6] Merchant classification codes were established decades ago and are not well suited to tracking newer categories of consumer spending. Many transactions at online grocery platforms carry codes for general or miscellaneous services, rather than the code for groceries.

precludes the use of publicly available datasets, such as from the Consumer Expenditure Survey (CEX), which follow a more limited number of households. The CEX also contains little information on online purchases beyond limited product categories and only began to include this information in more recent years. Other proprietary card processor transactions datasets have also become available in recent years. While these are also somewhat smaller in size, another important difference to the JPMC data is that the card is the anchor unit of analysis rather than the customer. Using customer identifiers in the JPMC data allows me to link transactions from multiple cards to one customer and the customer with their socioeconomic information as collected by lines of business. This detail allows me to study how spending behavior varies by key socioeconomic features.

2.A Customer Panel

I use the card transactions data to build a balanced panel for studying consumer trip and spending behavior over the sample window. To be included, customers must have a strong relationship with JPMC, meaning they make frequent purchases across a large number of products every month of the sample period.[7] This base customer panel contains approximately 7.7 million customers from across the US and largely matches the socioeconomics of the US population reported by the census. The sample skews somewhat male and high-income, reflecting that men are more likely to be listed as the primary account holder and more low-income consumers are unbanked. The sample is also more urban than the population as a whole, reflecting the footprint of the financial institution. Figure A1 shows the distribution of sex, age, and income among these customers.[8]

I focus in the analysis on goods and services which are well-represented on cards and, when purchased offline, typically involve customers and merchants in the same local market. They include: clothing, coffee, general goods, grocery, home maintenance goods and services, local leisure goods and services, personal care services, pharmacy, and restaurants.[9] Figure A2 summarizes the trip frequency and amount spent over time for customers in the panel. As with other datasets tracking consumer spending, groceries and restaurants dominate amongst everyday products. Spending patterns for all products exhibit strong seasonality and increases over time. The latter is likely driven by a combination of economic growth during the sample period and the increasing use of cards as payment instruments.

[7] The exclusion of new customers and customers that become inactive during my sample window accounts for the majority of the reduction from the base 69 million customer set. Many customers also use accounts exclusively for certain types of transactions (e.g. bill pay), rather than regular consumption. I exclude these and target banked customers who "home" on a their JPMC cards (Welander 2014; Cohen and Rysman 2013). Internal analysis comparing the card spend of strongly-tied customers to credit bureau data suggests the vast majority of their card spend is captured by JPMC data.

[8] Sex is imputed from names. Age is provided for customers 18 years and older. Yearly income is imputed from customer provided information and deposit account inflows. Customers with only credit card accounts are also treated as high-income because of credit account qualification requirements.

[9] General goods include department stores, discount stores, large non-specific online retailers, and other miscellaneous retailers like florists and books stores that sell everyday goods. Major categories of personal care services include salons and dry cleaners. Major categories of local leisure include movie theaters and gyms.

Patterns in the customer panel show broad scope for consumers to minimize the travel costs of their offline trips via both single-purchase trips and multi-purchase trips. To illustrate, Figure A3 Panel (a) shows that for the everyday products purchased most frequently by consumers, a large share of those purchases take place in the home zip code of the consumer. This includes groceries, for which 38% of offline transactions take place in consumers' home zip codes. This pattern suggests that a consumer's home is a common end point for trips including these products and that consumers minimize the travel costs for these products by purchasing them close to home.

On days when consumers shop offline, more than half of those days include multiple offline purchases. The distribution of the number of offline purchases is shown in Figure A3 Panel (b). On days with multiple purchases, consumers are likely to chain many of those purchases together on the same trip, saving on travel costs by both shopping close to home and at stores that are close to each other. For groceries, the most common products purchased on the same day are restaurants, general goods, and pharmacy, while for coffee, the most common pairings are restaurants, groceries, and general goods (Figure A4). The frequency of both sets of pairings is partly driven by overall product purchase frequency, but differences across groceries and coffee suggest that consumers have stronger preferences for some pairings over others.

While the primary focus of the analysis is the effect of online grocery platforms on consumers' offline shopping, I also include other products and channels in the panel to better understand the mechanisms that affect those offline purchases. I include online trips and spending for clothing, general goods, home goods and services, and restaurants. Figure A5 shows spending on major categories of online products over time as well as groceries. In addition, I include spending, regardless of channel, on fuel and local ground transportation.[10]

2.B Relevant Features of the Online Grocery Market

I leverage the entry of 17 online grocery platforms across different cities (Core Based Statistical Areas) for empirical identification of the effects of online grocery adoption. To pinpoint the month of entry of an online grocery platform into a city, I rely on a surge in transactions that are charged to that platform by customers who live in that city at the time of entry.[11] Using this measure of entry, these data show that more than 200 cities receive a first platform during the sample window and that multiple platforms often enter, with upwards of 10 platforms entering into the largest cities, like New York.[12]

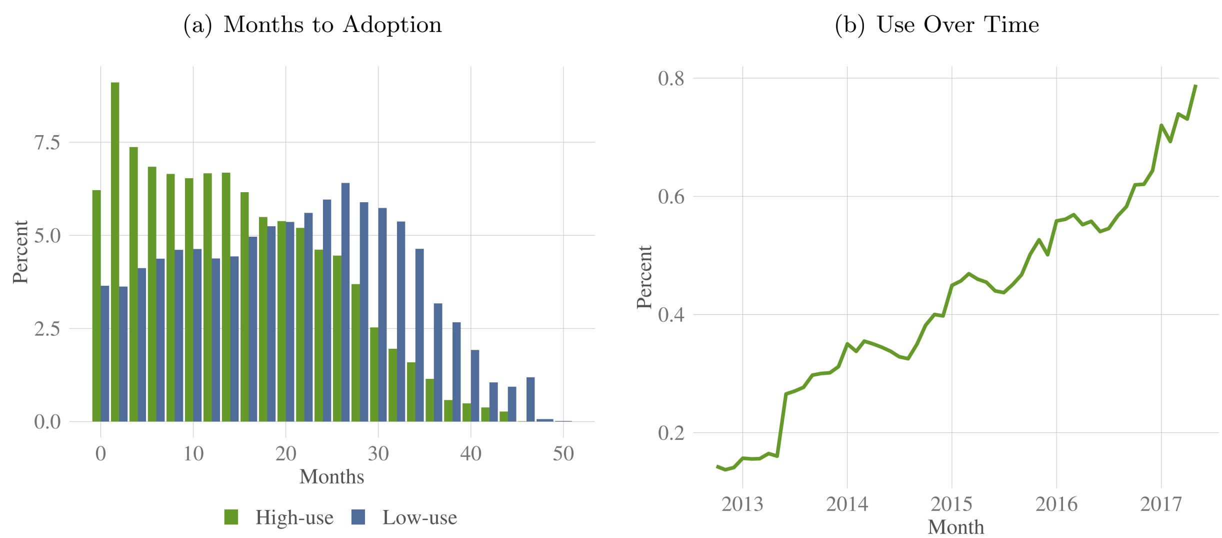
[10]Local ground transportation includes public transportation, taxis, parking lots, and ride-sharing services.

[11]Figure A6 Panel (a) shows that platform use at entry jumps from around 10 to more than 50 at the measured date of entry. Use prior to entry is due to mis-reported customer home zip codes and trial periods in some cities. The average masks wide variation; popular platforms in large cities have thousands of new customers at entry. Where possible, the entry measured by the card transactions was validated against publicly available data, including newspaper articles and company reports.

[12]Figure A6 Panel (b) shows the distribution of platforms per city at the start and end of the sample. Furthermore, Figure A7 Panel (a) shows an increasing pace of entry through 2015, suggesting most of the initial entry of platforms takes place over my sample window.

There are a number of features of the early online grocery market that make this entry exogenous to local consumer behavior. As with similar businesses with large fixed costs, online grocery platforms tended to target the most populous cities first, since a large customer base is key to profitability. They also targeted large cities near each other, reflecting economies of density (Holmes 2011). Furthermore, earlier platforms into a city had an advantage over later entrants in gaining market share, causing platforms to prioritize speed of entry. These factors are clearly reflected in the entry strategy for Amazon Fresh, for one example. After initial availability in Seattle in 2008, the platform quickly expanded to Los Angeles and San Francisco in 2013; San Diego, New York, and Philadelphia in 2014; Baltimore and Sacramento in 2015; and Boston, Dallas, and Chicago in 2016.[13]

Figure 1: Grocery Platform Use



Notes: Panel (a) shows the distribution of the number of months between the entry of a platform into a city and the first use of that platform by a user, split by low- and high-use. High-users use the platform in at least 5 separate months and are classified as adopters of the platform. A substantial share of high-use adopters begin using a platform in the first 12 months after entry. In contrast, low-users are more likely to first use a platform well after entry. There are 17 possible platforms for customers to adopt. Panel (b) shows the percent of customers using an online grocery platform in each month of the panel. A small percent of customers use the platforms, though the usage grows quickly.
Source: Panel (a) is calculated using the 77 and 26 thousand low- and high-users, respectively and Panel (b) is calculated using the 7.7 million customer panel.

Once available, there is wide variation in the adoption of online grocery platforms by customers in the city over time. The time between entry and adoption for customers who adopt a platform is seen in Figure 1 Panel (a). A significant share of customers who use the platform in at least 5 months (high-users) adopt an online grocery platform quickly after entry. However, many customers use the platforms for the first time much later, particularly those who do not use them on a continuing basis.[14] As a nascent online market, use of online grocery platforms across the panel population

[13] Entry dates from Wikipedia and not disclosed based on the card transaction data.

[14] About two-thirds of customers use a grocery platform for less than two months, but a quarter use them in a least 5 months. The distribution of use is shown in Figure A7 Panel (b).

is rare. Figure 1 Panel (b) shows that even though use of platforms grows quickly over time, less than 0.8% used one in the last month of my sample window. Adoption and use also surges in winter months, though platforms did not time their entry with this pattern – further evidence that platforms pursued quick entry during this period over targeting specific consumer consumption behaviors.[15]

3 Empirical Evidence

3.A Identification Strategy

The aim is to measure the effect of online grocery platform use on a consumer's offline shopping choices using a specification such as

$$Y_{izm} = \beta O_{izm} + \phi_m + \phi_{zq} + \mu_{izm}, \tag{1}$$

where Y_{izm} is an offline shopping outcome for consumer i in zip code z in month m, O_{izm} indicates use of the online grocery platform in month m, ϕ_m are month fixed effects, ϕ_{zq} are zip code by quarter fixed effects to control for different time trends in card spend across neighborhoods, and μ_{izm} is a consumer-specific error term.

There are two primary concerns in the measurement of the coefficient of interest, β. First, there are a number of endogenous choices on the demand and supply side of local markets that would cause the estimate of β to be biased due to correlations between consumer purchases and μ_{izm}. More specifically, consumers may experience shocks which change their preferred trips and stores might adjust their products and prices in response to market changes. Unobserved factors which affect these choices may make consumers more likely to make online purchases and change the composition of their offline purchases.

The exogenous entry of online grocery platforms across cities invites a standard difference-in-differences strategy for an unbiased estimate. One could measure consumption patterns in a city before and after platform entry against cities which have not yet experienced entry. However, the problem with this approach is statistical power. Even with large datasets, the infrequent use of platforms by customers in the early market makes β difficult to statistically significantly measure across the full city population (more so for spillover effects to non-grocery products). This necessitates an identification approach that leverages the exogenous entry of platforms, but focuses on consumers for whom platform entry affects their spending behavior.

To that end, I employ a modified difference-in-differences strategy. For the treatment group, I use a group of customers for whom the exogenous entry of the platform drives a plausibly exogenous

[15] Figure A7 Panel (b) shows the time series in platform entries and initial use of platforms by customers. There is no discernible pattern in entries that suggests they target winter entry dates.

shift in their online grocery purchasing behavior. These are "early adopters" – customers who become online grocery shoppers in the first 12 months after online grocery platform entry into their city and use the platforms for at least five months.[16] Although exact adoption timing and intensity is endogenous, I argue these early adopters would have likely used the platforms earlier if they had been available.[17] For them, unobserved endogenous changes that drive platform adoption are less likely to be correlated with the *timing* of their actual adoption decision. Therefore, any observed changes in spending patterns at the time of platform adoption can be attributed to the effect of the platform itself, rather than those unobserved factors.

Of course, early adopters of online grocery platforms systematically differ from the general population, as Table B2 shows. Early adopters are more likely to be female, younger, higher income, and spend more on more trips than non-adopters. They also spend substantially more online on non-grocery products and travel more via local ground transportation prior to adoption. Therefore, I create a control group of similar, plausible adopters who, for some reason, chose not to adopt a platform during my sample window. To be a valid approach, it must be the case that the matched non-users do not adopt a platform for reasons exogenous to their shopping behaviors. These possible exogenous reasons include differential exposure to early advertising and slow learning about the new platforms through social networks (Butters 1977; Conley and Udry 2010; Goolsbee and Klenow 2002; Bell and Son 2007).[18]

To create these controls, I match each early adopter to two platform non-adopters from the same set of zip codes using nearest neighbor matching on pre-adoption socioeconomics and spending patterns.[19] Matching is done separately for each month based on the month of platform adoption. The specification is

$$EA_{im} = \gamma^1 Sex_i + \gamma^2 Age_{im-1} + \gamma^3 Income_{im-1} + X_{im-1} + \nu_{im}, \tag{2}$$

where EA_{im} indicates that customer i is an early adopter in month m, Sex_i, Age_{im-1}, and $Income_{im-1}$, are vectors indicating the customers' sex, pre-adoption age group, and pre-adoption income group, respectively, and X_{im-1} is a vector capturing levels and changes in spending and trip patterns in the year prior to adoption.[20] I include changes over three- and 12-month intervals

[16] I choose 12 months to capture those who adopt immediately post-entry and who adopt the first winter after entry, given the strong relationship between season and adoption that I observe in the data. I choose 5 months as a cutoff to qualify for adoption to avoid customers who only use platforms for a limited time because of dissatisfaction or temporary promotions.

[17] More generally, one might also be concerned with more long-term endogenous choices, such as consumers who sort to live near their preferred stores and stores which enter and exit based on location profitability. These are unlikely to be at play in the short window around platform entry and adoption.

[18] Customers who only briefly try a platform do not make a good control group because their spending does not return to pre-adoption levels. This may be due to subsequent experimentation with other food-delivery options.

[19] This is similar in approach to new difference-in-differences estimators based on the work of Callaway and Sant'Anna (2021) and Sun and Abraham (2021), in which propensity score matching can be utilized to improve parallel trends between treatments and controls. I match as a separate step because I use a random sample of potential controls due to the large sample size.

[20] I include a subset of offline products (restaurant, grocery, leisure, pharmacy, personal, and coffee) and all online

in X_{im-1} to capture both seasonal variation and differential time trends in card spending.[21] Using these short and long changes obviates the need for month or zip by quarter fixed effects, the ϕ_m and ϕ_{zq} in equation 1, when deploying the same matched data in the later structural estimation.

My difference-in-differences strategy then compares how the purchase decisions of early adopters change in the immediate window around the exogenous timing of their adoption decision compared to a set of non-users who were similar, plausible adopters. Denote G_i as the month of grocery adoption for customer i and $D^l_{im} = \mathbb{1}\{t - G_i = l\}$ as the relative month l of adoption dummy for customer i. I use the specification

$$Y_{im} = \sum_{l=-K, l\neq -1}^{L} \delta_l D^l_{im} \times EA_i + \mu_{im}, \tag{3}$$

where the δ_l for early adopters measures the effect from adoption of an online grocery platform at month l relative to their matched controls. For more detailed choices that occur at lower frequencies, I measure the average change in the post-adoption period with the specification

$$Y_{im} = \delta_0 D_{im} + \delta_1 EA_i + \delta_2 D_{im} \times EA_i + \mu_{icm}. \tag{4}$$

where $D_{im} = \mathbb{1}\{t - G_i \geq 0\}$.

I also note that, although early adopters of online grocery platforms are a selected sample, the aim of this research is not to recover a set of effects common to the full population. The aim is show that products can be winners, not just losers, to online retail and explore mechanisms that can create them. Those mechanisms are unlikely to be specific to this sample, even if the exact effects are sample-specific.

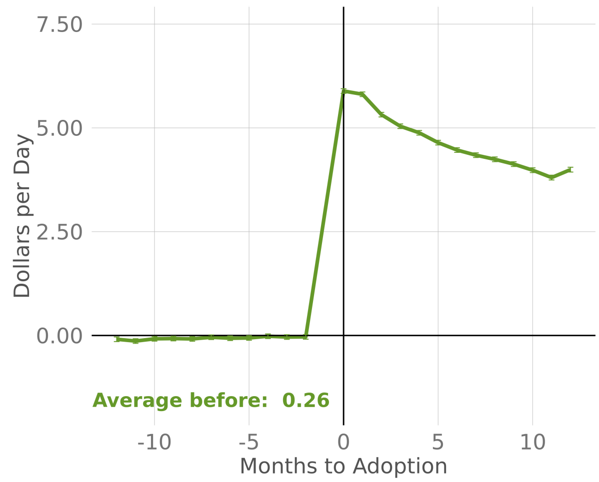
3.B Changes in Grocery Purchases

Using this empirical strategy, I find that early adopters of online grocery platforms spend at a rate of \$5.89 per day on online grocery platforms in the month of adoption. This is shown in Figure 2. Over subsequent months, spending moderates as many users do not use the platforms each month and some stop using the platform altogether. Still, over the subsequent year, online grocery spending remains elevated over \$3.80 per day and even increases at the end of 12 months when many consumers experience the start of another winter season.

products as well as fuel and transportation in the matching estimation. For grocery and coffee trips, I match on coffee and grocery trip bundling because I focus on finer-grained trip patterns for these products.

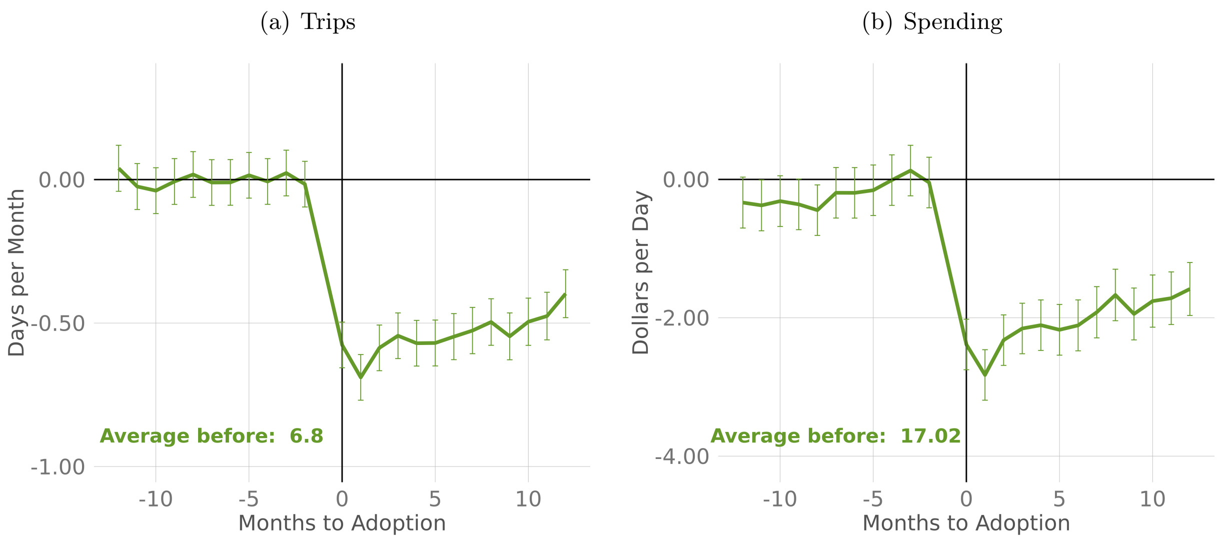
[21] Table B1 shows the coefficients for equation 2 for adoption in October 2014. Tables B2 and B3 show improvements in two-sided t-stats after matching for covariates used in the matching procedure. Reassuringly, Table B4 also shows similar improvements on offline product categories not used in the matching procedure, including clothing, home goods and personal care services.

Figure 2: Online Grocery Spending



Notes: This figure shows that change in offline and online spending on groceries the 12 months before and after platform adoption for early adopters of platforms as compared to a matched sample of non-users. This figure, as well as subsequent figures, also displays the average number of days for grocery trips and dollars spent per day in the month prior to adoption.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on pre-adoption demographics and spending patterns.

Figure 3: Grocery Store Trips and Spending



Notes: This figure shows the change in trips at offline grocery stores in the 12 months before and after platform adoption for early adopters of platforms as compared to a matched sample of non-users. In the month of adoption, consumers reduce their daily trips by 0.6 days per month and by 6.4 trips over the following 12 months. This is a 7.9 percent decline from the month prior to adoption, on average. In terms of spending, consumers reduce their daily spending at grocery stores by an average of 12.0 percent. Early adopters increase their spending on groceries overall after adoption by $2.61, or 15.1 percent.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on zip code, socioeconomics, and pre-adoption spending patterns.

In response, I find that early adopters of online grocery platforms significantly reduce their trips and spending at grocery stores after adoption. Figure 3 shows that prior to adoption, early adopters made trips to grocery stores 6.8 days per month and spent \$17.02 per day at the stores.[22] In the month of adoption, consumers reduce their daily trips by 0.6 days, and over the following 12 months make a total of 6.4 fewer trips (an average decline of 0.5 trips per month). On average, this is a 7.9 percent decline from the month prior to adoption. In terms of spending, consumers reduce their daily spending at grocery stores by \$2.05 per day in the post period, on average, a reduction of 12.0 percent. This reduction matches the classic substitution effect for direct competitors to online retail. Finally, note that overall grocery spending after platform adoption increases by 15.1 percent. This could be driven partly by more overall grocery consumption, online price differences (as online groceries tend to be more expensive), and any included delivery service charges.

3.C Changes in Trips Across Other Products

I find that the adoption of an online grocery platform sharply changes consumption across a wide selection of goods and services in addition to groceries. Figures 4 and 5 show the effects on offline trips for each product set, respectively (effects in terms of spending for goods and services are qualitatively similar). For goods, only pharmacy trips are strongly affected with an increase equivalent to 0.5 additional trips over 12 months. In contrast, every local service increases. For coffee, local leisure, personal care, and restaurants, trips increase at a rate of 1.0, 0.3, 0.9, and 2.6 extra trips, respectively, over 12 months after platform adoption.[23] Further evidence that these are additional trips can be seen in the changes to travel-related spending. Figure 6 shows a 12.9 average increase in travel spending via local ground transportation after platform adoption.[24] Of course, the broader set of new trips reflected in this spending could include both retail trips and trips without card spending (e.g. visits to friends and family).

There are a number of mechanisms that could drive offline trip increases in non-competing products. These mechanisms include income effects, consumption complementarity, and online and offline shopping complementarities. In the case of online groceries, income effects are unlikely to increase new trips for other products. Given that consumers who use online grocery platforms spend more on groceries overall after platform adoption, one would expect that any income effects would reduce, rather than increase, other offline trips.

In consumption complementarity, the joint consumption of products provides higher utility

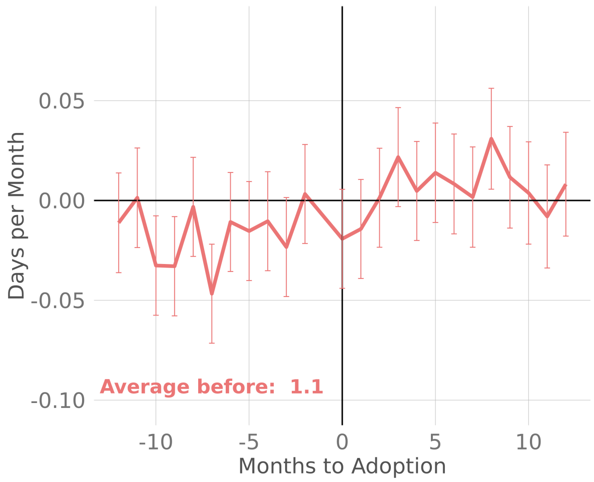
[22] This closely matches a 2019 industry report on grocery shopping trends which found that the primary grocery shopper for a household made around 1.5 grocery store trips per week during my sample period.

[23] Results for late adopters of online grocery platforms (those who use platforms at least 12 months after entry and for at least 5 months) are broadly similar. For example, Figure B1 shows that they reduce their grocery store trip frequency by slightly more and increase their coffee trip frequency slightly less than their matched controls.

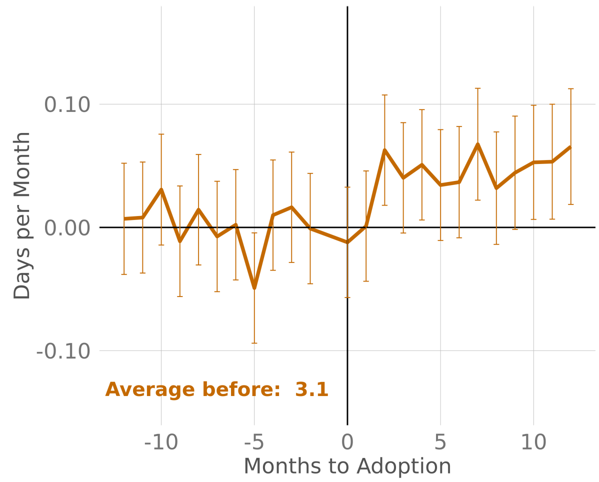
[24] I show travel-related spending rather than trips to abstract from customers who load large amounts of money on transportation cards. Although not statistically significant, the results also suggest some reduction in fuel spending. This makes sense if grocery trips are more likely to be made by car to accommodate heavy grocery bags while service consumption is comparatively easy via public transportation.

Figure 4: Local Goods Trips

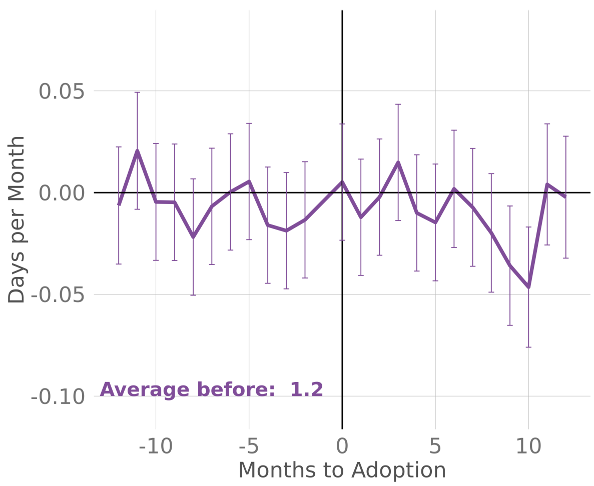
(a) Clothing



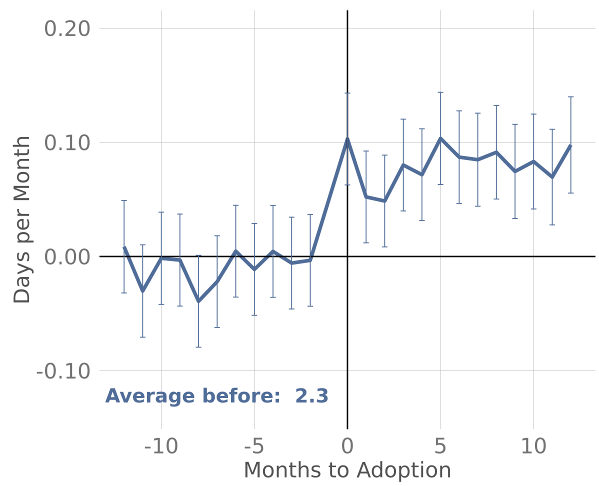
(b) General Goods



(c) Home Goods



(d) Pharmacy

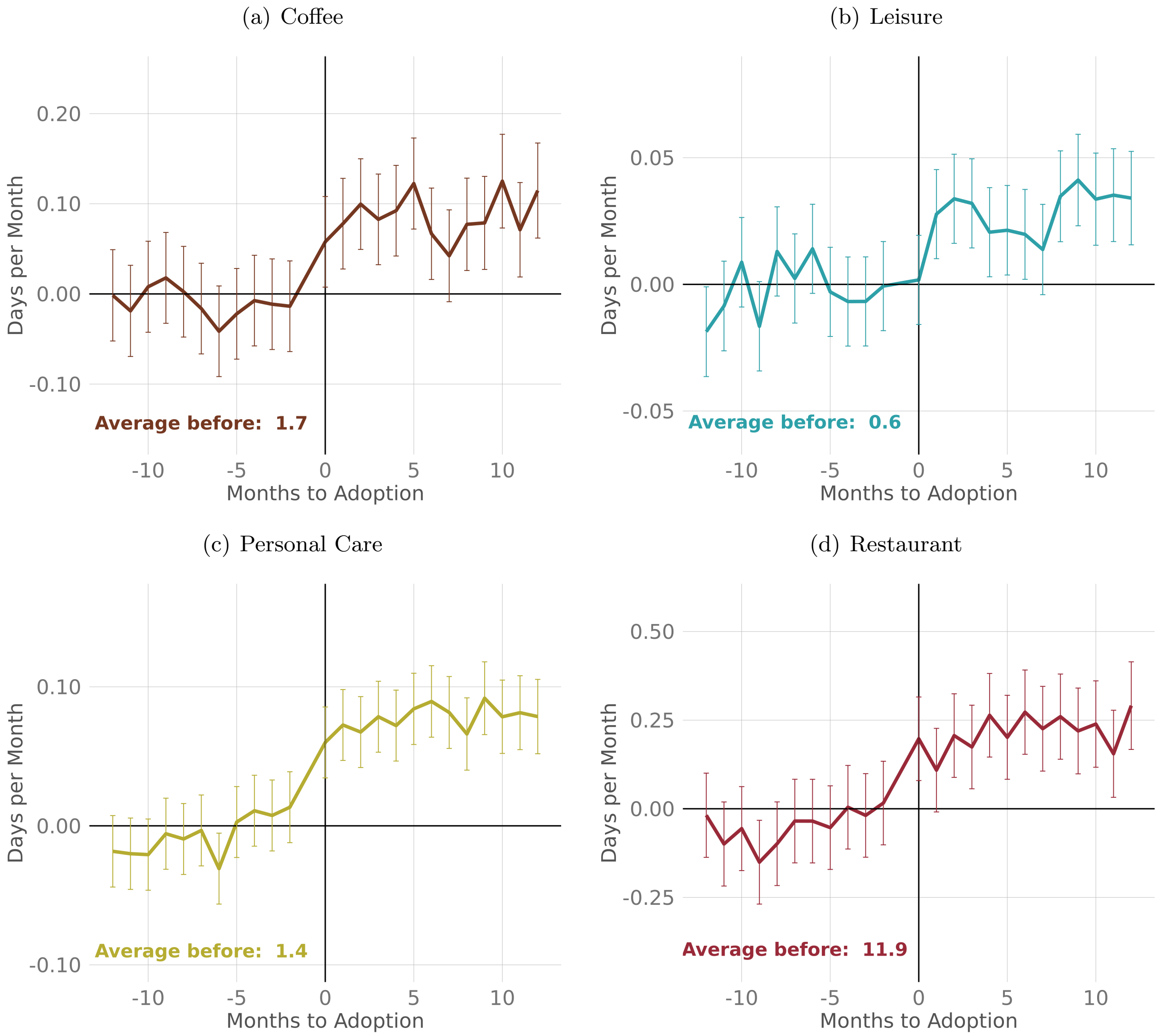


Notes: This figure shows the change in the frequency of trips including local goods in the 12 months before and after platform adoption for early adopters of platforms as compared to a matched sample of non-users. General goods include department stores, discount stores, large non-specific online retailers, and other miscellaneous retailers like florists and books stores that sell everyday goods. In the 12 months after adoption, early adopters largely maintain their spending on local goods, except for pharmacy, for which customers increase their trip frequency at a rate of 0.5 days over a year. The figure also displays the average number of days for each trip in the month prior to adoption.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on zip code, socioeconomics, and pre-adoption spending patterns.

than the equivalent consumption of each good alone, as with pasta and pasta sauce (Mehta and Ma 2012). The increase in pharmacies and restaurants is most likely to be partly explained by this complementarity with online groceries. In the case of pharmacies, they may act as convenient corner stores to top-up food purchases and purchase over-the-counter medications for consumers who use online grocery platforms.[25] In the case of restaurants, online grocery users may have a

[25] Medication which is purchased at grocery stores' internal pharmacies are classified as pharmacy purchases because

Figure 5: Local Services Trips



Notes: This figure shows the change in the frequency of trips including local services in the 12 months before and after platform adoption for early adopters of platforms as compared to a matched sample of non-users. Major categories of personal care services include salons and dry cleaners. Major categories of local leisure include movie theaters and gyms. In the 12 months after adoption, early adopters significantly increase their spending on local services. Coffee, leisure, personal care, and restaurant trips increase at a rate of 1.0, 0.3, 0.9, and 2.6 days more over a year, respectively. The figure also displays the average number of days for each trip in the month prior to adoption.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on zip code, socioeconomics, and pre-adoption spending patterns.

higher demand for prepared food after platform adoption if they tend to use platforms for packaged and staple foods, as suggested by surveys.[26]

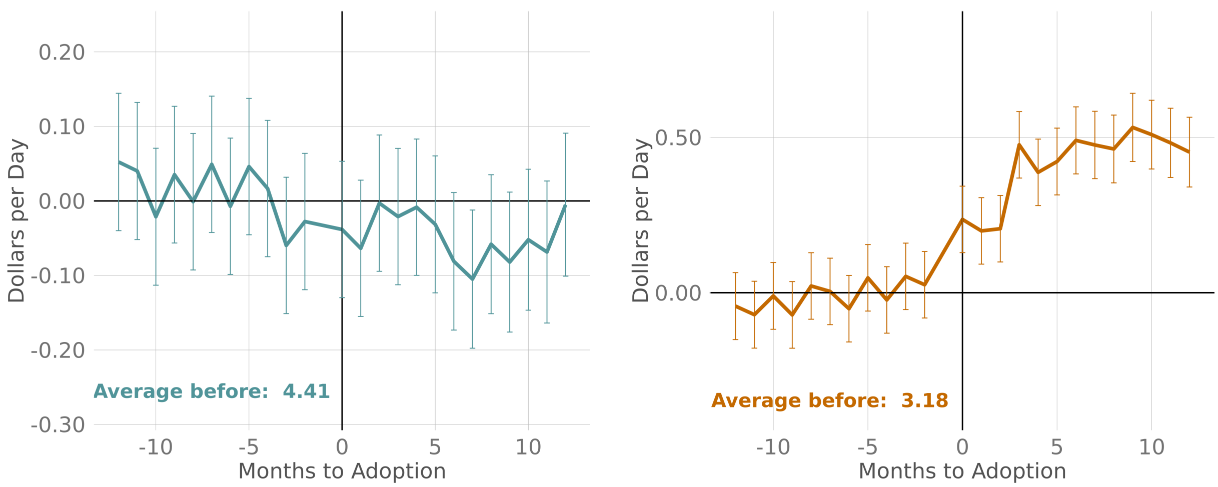
Shopping complementarities, across online and offline channels, may also play a role (Baker et

they typically have distinct card terminals. Therefore, this increase is less likely to directly reflect changes in where prescriptions are filled. Though changing where the prescription is filled could have spillover effects to other pharmacy purchases.

[26]See U.S. Grocery Shopper Trends for 2019 for one.

Figure 6: Local Travel Spending

(a) Fuel Spend (b) Transportation Spend



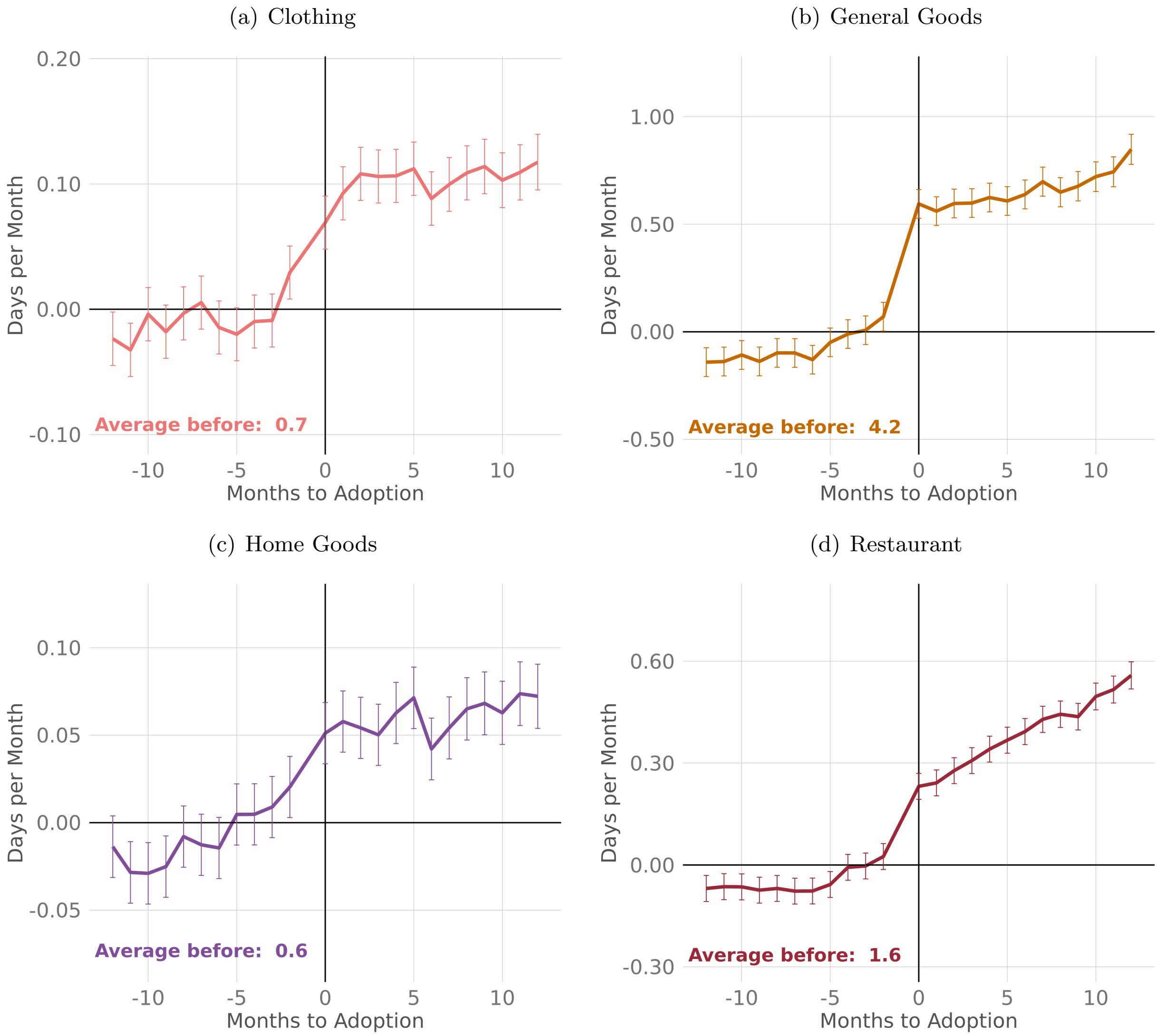
Notes: This figure shows the change in spending on local travel-related expenses in the 12 months before and after platform adoption for early adopters of platforms as compared to a matched sample of non-users. After adoption, early adopters significantly increase their spending on local ground transportation (including taxis, public transportation, and ride-sharing) by an average of $0.41 per day, or 12.9 percent. The figure also displays the average dollars per day spent on each travel-related expense prior to adoption.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on zip code, socioeconomics, and pre-adoption spending patterns.

al. 2020). For instance, when consumers choose to purchase groceries online, they may be more likely to chain virtual visits within or across online retailers to purchase additional products. Figure 7 shows that consumers do indeed concurrently increase trips for other major online products after grocery platform adoption. However, such increased consumption online would be expected, if anything, to decrease trips for these products offline through direct substitution. Instead, most offline equivalents show no change or an increase.

Rather, the most likely mechanism to consistently favor increases in services over goods is an offline shopping complementarity created through changes in time use. For instance, for each grocery trip not taken, consumers can use that time saved to make new trips. An analysis of the American Time Use Survey by the USDA suggests that, around the start of my sample period, the national average for time used on a grocery store trip was approximately 75 minutes, including time spent traveling to the grocery store and on in-store shopping.[27] Combined with the result of 0.6 fewer grocery trips per month, that would imply 44 extra minutes every month to do something else. Consumers may be more likely to increase their frequency of service trips with that extra time, given that such trips are both discretionary and time-intensive.

[27]See How Much Time Do Americans Spend on Food? and Access to Affordable and Nutritious Food. In addition, the 2019 American Time Use Survey shows that consumers spent around 5.2 hours per week purchasing all goods and services (including in-store and online purchasing), around an hour of which was travel-related. This gives a sense of the upper bound of time-savings that could be realized from further online shopping.

Figure 7: Online Trips



Notes: This figure shows the change in trips for other online products in the 12 months before and after platform adoption for early adopters of platforms as compared to a matched sample of non-users. In the months after adoption of an online grocery platform, early adopters increase their online trips for other online products.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on pre-adoption demographics and spending patterns.

In the case of coffee, local leisure, and personal care services, this offline shopping complementarity through time use is likely the primary mechanism that increases trips. These services have no obvious consumption complementarity with groceries and are difficult, if not impossible, to replicate with online versions. For example, assuming online grocery platform use has no impact on home coffee drinking habits, there is no clear reason why consumers would change their taste for coffee from coffee shops in response to platform use. Furthermore, hot, fresh-brewed coffee and the coffee shop experience are not easily delivered via the online channel.[28]

[28]The partnership between Starbucks and Uber Eats occurs after my sample period.

3.D Evidence from Bundled Trips to Grocery Stores and Coffee Shops

To look more closely at the impact of time use on consumers' trip choices, I focus on the effects of online grocery platform use on bundled trips to grocery stores and coffee shops. Limiting the focus to two product types maintains tractability in a setting where combinatorics quickly multiplies the choice set while still highlighting the focal mechanisms. I focus on coffee shops because they are one of the services I track which is unlikely to be affected by alternative mechanisms that can induce consumer substitution when consumers purchase more groceries online. Coffee shops are also attractive because consumers transact at them relatively frequently and because they are fairly uniform and numerous, providing ample substitution opportunities for consumers across space in their offline trip choices. I emphasize, however, that coffee shops are used for illustration and that any product could be affected by this same mechanism.

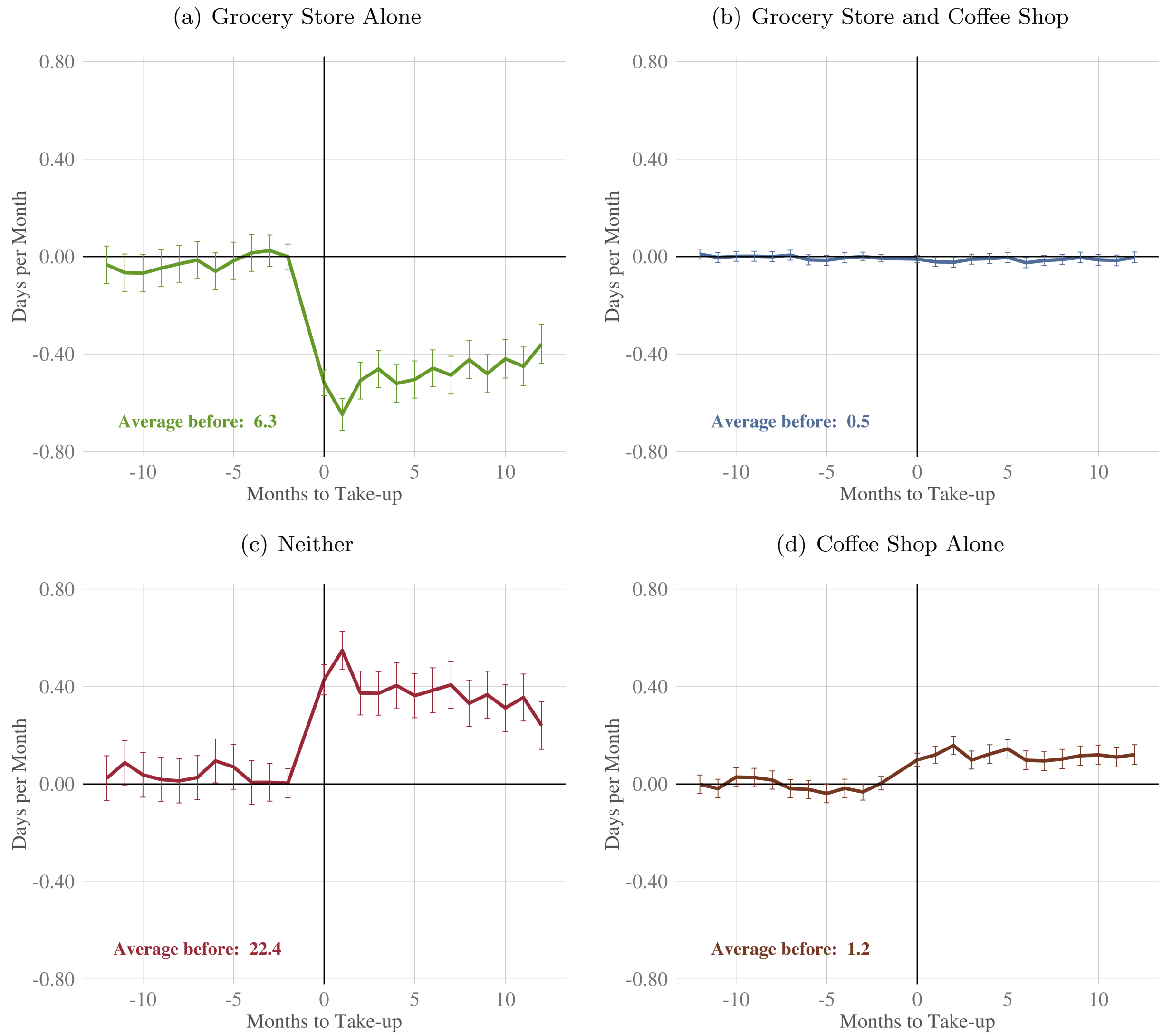
I make a few trivial assumptions on trip formation which I later carry over into the discrete choice model. I assume that groceries and coffee are consumed every day so that consumers must make a trip each day with non-zero cost. However, groceries are durable and do not need to be procured every day, while hot, fresh-brewed coffee is not durable and needs to either be purchased or made at home every day.[29] This results in four possible trip types for a consumer who makes exactly one offline shopping trip each day: (1) grocery store alone, (2) coffee shop alone, (3) grocery store and coffee shop, and (4) neither store with coffee made at home.

Using this two-product setup, I find that consumers are not just replacing trips to the grocery store to go on new trips, but changing how they make single- and multi-store trips together across time and space. Figure 8 shows the average change in the four bundled trip types. Most of consumers' trip substitution after platform adoption is from trips only to the grocery store toward neither store. More interestingly, however, is that consumers decreased their combined trips to both the grocery store and coffee shop such that the effect for coffee shops from the previous section is likely driven by two separate substitutions: one from a trip only to the grocery store to a trip only to the coffee shop and one from the chained trip to both stores to the trip only to the coffee shop.

Patterns in trip substitution across days of the week after platform adoption support that different opportunity costs of time affect trip choices (Figure 9). One interesting pattern is for Sundays, when consumers were most likely to make a trip to the grocery store (with and without the coffee shop) before platform adoption. I find that on Sundays after platform adoption, consumers reduce their trips which include the grocery store the most on Sundays, primarily substituting toward trips to neither store that day. This may reflect that on Sundays consumers have a high preference for non-shopping activities, such as time with friends and family. Another pattern that emerges is the stronger work-week versus weekend increase for trips only to the coffee shop. Consumers who work on weekdays might prefer to consume more coffee on those days, perhaps for

[29] The 2017 National Coffee Association survey estimated that 62 percent of adults drank coffee in the previous 24 hours and over 80 percent made coffee at home.

Figure 8: Bundled Trips

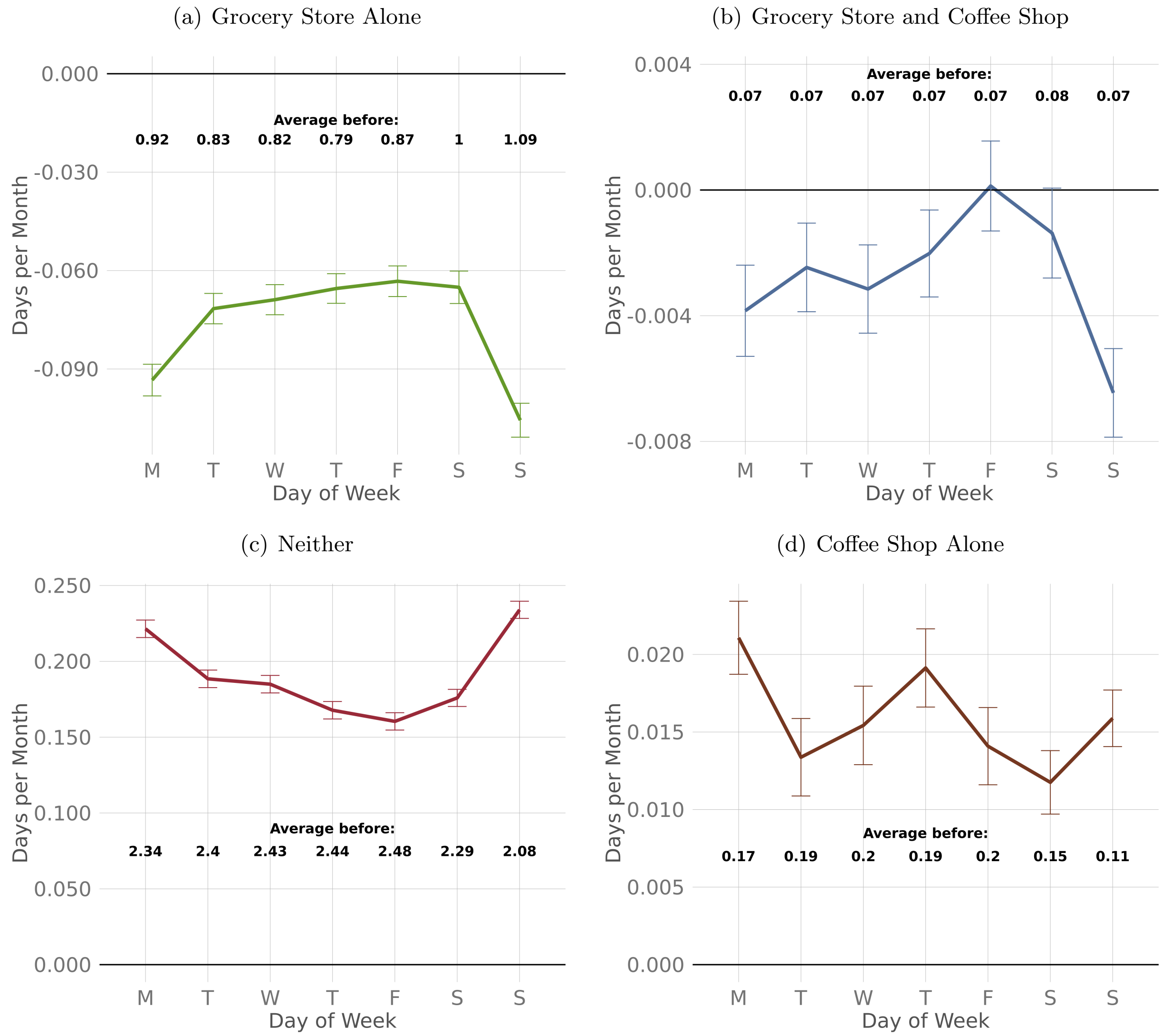


Notes: This figure shows that early adopters change the composition of their bundled trips to grocery stores and coffee shops after they adoption of an online grocery platform. The figure also displays the average number of days for each trip in the month prior to adoption. The dominant effect is to shift trips only to a grocery store to trips to neither store. However, at the same time, consumers decrease their trips to both grocery stores and coffee shops and increase trips only to the coffee shop.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on zip code, socioeconomics, and pre-adoption spending patterns.

stimulation or to socialize with colleagues.

In addition, I find that the consumers reorganize toward new single- and multi-store trip chains with lower distance costs. Figure 10 Panel (a) shows that on trips only to the coffee shop, the post platform adoption increase is disproportionately toward coffee shops located close to where the consumer lives. And Panel (b) shows that for trips to both the grocery store and the coffee shop, combined trips in which the two stores are closer to each other are more common post adoption

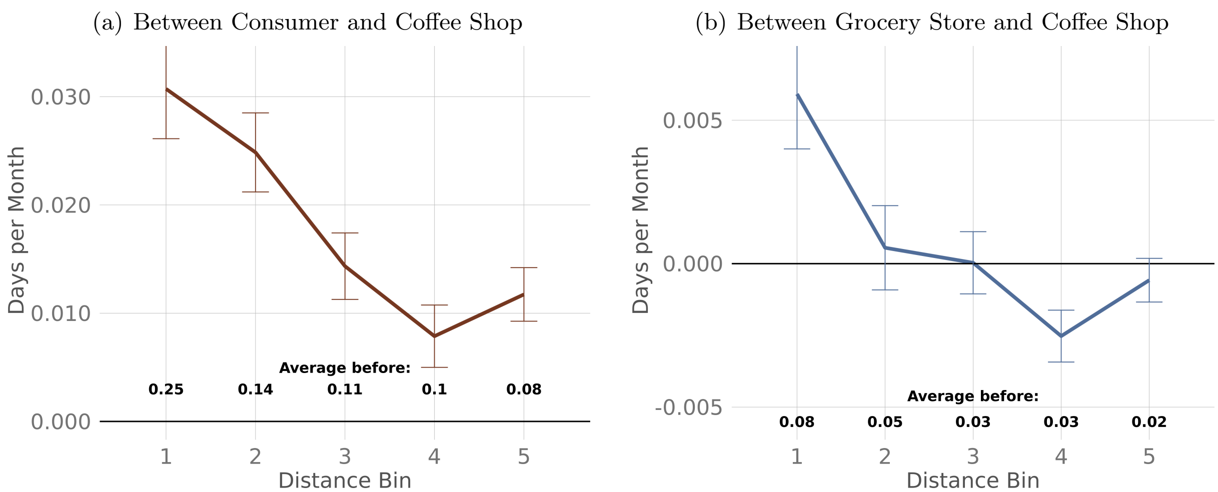
Figure 9: Bundled Trips by Day of Week



Notes: This figure shows the average post-adoption effect by day of week for each bundled trip type. The results highlight the importance of consumers' daily time budgets to the composition of their trips to grocery stores and coffee shops after they adopt an online grocery platform.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on zip code, socioeconomics, and pre-adoption spending patterns.

even though combined trips overall are less likely. Thus, coffee shops closer to consumers win more and grocery stores close to coffee shops lose less from the negative effects of platform adoption. However, in locations without such accessibility to winning stores like coffee shops, consumers may instead substitute toward more non-shopping activities. These results suggest that time use is key to creating offline winners to online retail.

Figure 10: Bundled Trips by Selected Distances



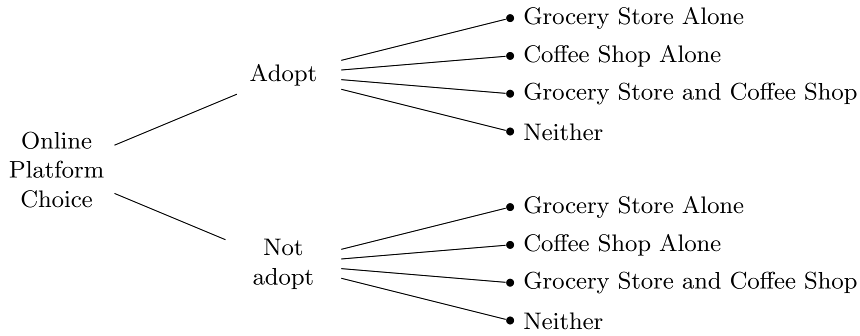
Notes: Panel (a) shows the average post-adoption effect for the coffee alone trip by coffee shop distance. Panel (b) shows the average post-adoption effect for the combined grocery store and coffee shop trip by the distance between the grocery store and coffee shop. Distances are measured between zip code centroids. Distance bins increase from left to right, with the first distance bin implying the same zip code. Distance bins 2-5 imply distances from 2-5 miles, respectively. There is a clear distance gradient, with consumers increasing their trips to coffee shops alone disproportionately to nearby coffee shops and increasing their trips to both grocery store and coffee shops when those stores are close to each other.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and each of their two nearest neighbors matched on zip code, socioeconomics, and pre-adoption spending patterns.

4 A Trip Choice Model with Flexible Substitution Patterns

I now lay out a discrete choice model of consumer shopping trip choice which replicates the rich substitution patterns evident in the previous empirical exercise. As laid out in the two-stage decision tree in Figure 11, I model a consumer who first makes a long-run decision to adopt an online grocery platform or not and then, conditional on that choice, which shopping trip to make each day. To maintain tractability, I continue in the model with the four trip options over groceries and coffee and further assume that the consumer only has one grocery store, g, and one coffee shop, c, to choose from on one shopping trip per day. All consumers, indexed by i, are ex-ante identical apart from their location and income. The model is solved by backward induction, with the consumer first determining the expected value of her daily trip choices conditional on her long-run online grocery shopping decision. I assume a fixed supply-side and no general equilibrium effects.[30]

[30]The low rates use of online grocery platforms during this period imply that, in the short-run, the availability of new online groceries was a marginal change in the market and unlikely to induce wide general equilibrium effects on the demand and supply of goods and services across locations. These would include, among others, consumer location choices, store entry and exit decisions, and the pricing of goods and services. This implies that estimates from this study are best interpreted as partial equilibrium effects conditional on prevailing market conditions.

Figure 11: Consumer Decision Tree



Notes: This figure shows the consumer's two-level discrete choice problem for selecting groceries and coffee. In the long-run, they decide whether or not to be an online grocery shopper. Then, conditional on that choice, they decide which of four possible offline shopping trips to make each day.

4.A Daily Trip Choice

I assume consumer i's utility maximization problem for choosing one of the four trips on a day takes the form

$$\max_{g,c} V_i(g,c) = \beta_0(g,c) + \tau \ln(d_i(g,c)) + \epsilon_i(g,c), \tag{5}$$

where $V_i(g,c)$ is the trip value based on the store set $\{g,c\}$, with $g=1$ and $c=1$ signifying the inclusion of the grocery store or coffee shop, respectively. Values for the stores included on a trip are defined by

$$\beta_0(g,c) = G\mathbb{1}_{\{g=1\}} + C\mathbb{1}_{\{c=1\}} + b\mathbb{1}_{\{g=1,c=1\}}, \tag{6}$$

where G is the grocery store value, C is the coffee shop value, and b is a fixed benefit shifter for a trip that includes the grocery store and coffee shop. The fixed benefit for grocery and coffee is likely to primarily reflect non-distance benefits of trip-chaining (e.g. leaving home once), but for other goods could also reflect other complementarities in joint consumption as in Gentzkow (2007). Trip utility is reduced by τ, the opportunity cost of time, times the log distance, $d_i(g,c)$, traveled on each trip in miles. The four trip distance costs are (1) $d_i(1,0) = 2d_i^g$, (2) $d_i(0,1) = 2d_i^c$, (3) $d_i(1,1) = d_i^g + d_i^c + d^b$, and (4) $d_i(0,0) = 0.1$, where d_i^g is the distance to the grocery store from the consumer's home (likewise for d_i^c), and d^b is the distance between the grocery store and coffee shop. I assume that $|d_i^g - d_i^c| < b + d^b \leq d_i^g + d_i^c$, so that travel to both stores on one trip is always more costly than travel to just one store and travel to both stores on one trip is less costly than visiting each store separately on the same day. The cost of the trip to neither store, with coffee at home, is equivalent to 0.1 miles.[31] In addition, the consumer has a random taste shock, $\epsilon_i(g,c)$, for each trip.

[31] This is a practical assumption to avoid taking the log of zero for trip distance costs, but also reflects that producing coffee at home has its own time cost.

Consumers are separated into low- and high-income groups with separate trip utility parameters estimated for each group to capture different store values and opportunity costs of time.[32]

The central challenge of this setting is the correlation across the values of trips for an individual consumer in a day. One way these can be created is through the fact that the value of a store in a trip will affect the values of other trips that also include that store. For example, on a day when a consumer wakes up wanting coffee from the coffee shop, any trip containing a visit to the coffee shop should be more likely for her, leading to a positive correlation between $\epsilon(0, 1)$ and $\epsilon(1, 1)$. Another way these can be created is through shocks to a consumer's time budget. For example, on a day when consumers have little time for shopping, trips with short distance costs should be more likely. Therefore, models which carry independence of irrelevant alternatives (IIA) assumptions assume away the correlations that are fundamental to trip substitution patterns. These include the classic logit and nested logit models as well as BLP-type models based on Berry et al. (1995). This is similar to the argument in Compiani (2021), who uses simulations to show that BLP performs poorly in settings with complementarities.

Another challenge to modeling trip substitution is that consumers face trip distance costs specific to their individual locations.[33] This is distinct from the typical discrete choice problem in which consumers in the same market face the same menu of costs. To illustrate, imagine a world with one grocery store and one coffee shop and two identical consumers, A and B, except that A is closer to the coffee shop and B is closer to the grocery store.[34] These different relative distances to the stores imply that any shock to the value of a trip should impact them differently. For example, imagine that A and B each wake up one day with the same craving for coffee from the coffee shop, leading to high $\epsilon_i(g, c)$ shocks for both consumers for trips that include the coffee shop. But, because consumer B is farther from the coffee shop, we could expect her to be less likely to make a trip including the coffee shop, even with the craving. Or if she does make the trip, the coffee shop's proximity to the grocery store might make her more likely to visit both the coffee shop and the grocery store to economize on travel costs. However, consumer A might be much more likely to visit the coffee shop alone since she has lower travel costs for that trip type.

Therefore, this context requires a model that accounts for the full set of cross elasticities among multiple choices with individual-specific distance costs. To that end, I use the paired combinatorial logit model (PCL) from the transportation literature, which explicitly models the choice elasticity for each pair of choices. The properties of this model and its relation to other discrete choice models with type I extreme value errors are described in Koppelman and Wen (2000).

In a slight abuse of notation for concision, denote the four possible trip choices $\{g, c\}$ as

[32] Consumers can be classified as high-income in two ways. First, their income in 2013 is more than \$50,000. Second, they can be customers with only a credit card account. For the latter, the bank does not provide an income estimate, but such customers tend to skew higher-income.

[33] This distinction has been made before in some studies of retail markets, such as that for gasoline in Houde (2012) and grocery stores in Thomassen et al. (2017).

[34] Figure C2 shows this setup in a simple diagram.

$gc \ \epsilon \ \{10, 11, 01, 00\}$. In the model, the probability of trip gc depends on the probability of trip gc relative to trip gc', $P_{gc|gc,gc'}$, and the value of a pair of trips relative to other pairs, $P_{gc,gc'}$, across the possible $gc \neq gc'$. The total probability of a trip gc sums over the product of these two terms for all possible trips paired with gc,

$$P_{gc} = \sum_{gc' \neq gc} P_{gc|gc,gc'} P_{gc,gc'}, \tag{7}$$

where

$$P_{gc|gc,gc'} = \frac{\exp\left(\frac{V_{gc}}{1-\sigma_{gc,gc'}}\right)}{\exp\left(\frac{V_{gc}}{1-\sigma_{gc,gc'}}\right) + \exp\left(\frac{V_{gc}}{1-\sigma_{gc,gc'}}\right)},$$

$$P_{gc,gc'} = \frac{\exp((1-\sigma_{gc,gc'})I_{gc,gc'})}{\sum_{ij}\sum_{ij' \neq ij} \exp((1-\sigma_n)I_{ij,ij'})},$$

and

$$I_{gc,gc'} = \ln\Big[\exp\Big(\frac{V_{gc}}{1-\sigma_{gc,gc'}}\Big) + \exp\Big(\frac{V_{gc'}}{1-\sigma_{gc,gc'}}\Big)\Big].$$

$I_{gc,gc'}$ is the inclusive value of trip pair gc, gc' and $0 < \sigma_{gc,gc'} < 1$ is a similarity index that captures the substitutability between the two trips, with $\sigma_{gc,gc'} = 1$ for perfect substitutes. Because there are six possible pairs of trips, there are six substitutability parameters, denoted $\boldsymbol{\sigma}$. This model has a similar structure to the traditional logit and nested logit models. It uses a logit shock over the value of a trip within a trip pair combined with a logit shock for the value of the trip pair. This allows simultaneous nesting of any trip with every other trip to account for different degrees of substitution between each pair of trips. As a result, trip gc is more likely if trip gc is valuable relative to trip gc' or the value of the pair of trips gc and gc' is high (due to high trip utility parameters in the two trips or high similarity between them).

4.B Substitutions in Trip Choice

To illustrate the rich patterns of substitution captured in the model, consider a decline in the value of the grocery store and its effect on the relative attractiveness of two trips. For example the log relative probability for only coffee versus the neither trip is

$$\ln\Big(\frac{P_{01}}{P_{00}}\Big) = \ln \sum_{gc' \neq 01} V_{01|01,gc'} - \ln \sum_{gc' \neq 00} V_{00|00,gc'}, \tag{8}$$

where

$$V_{gc|gc,gc'} = \exp\Big(\frac{V_{gc}}{1-\sigma_{gc,gc'}} - \sigma_{gc,gc'} I_{gc,gc'}\Big)$$

acts as a pair-weighted trip value. Through the trip pair inclusive values, $I_{gc,gc'}$, in each of the $V_{gc|gc,gc'}$ terms, the elasticities of the relative probability of the two trips are functions of trip utility parameters that are unrelated to the coffee alone or neither trip. This includes grocery store value, G, but also the distances to the grocery store and between stores, d^g and d^b, and the fixed benefit to the chained trip, b. The effects are then scaled by the strength of consumers substitution between trips in each pair, $\boldsymbol{\sigma}$, and the utility cost of distance, τ.

Therefore, when the value of the grocery store falls, the value of the coffee alone trip adjusted by the weight of the chained trip and coffee alone trip pair increases,

$$-\frac{\partial V_{01|01,11}}{\partial G} = \frac{\sigma_{01,11}}{1-\sigma_{01,11}} P_{11|01,11} V_{01|01,11} > 0. \tag{9}$$

Similarly, the value of the coffee alone trip adjusted by the weight of the grocery alone and coffee alone trip pair increases,

$$-\frac{\partial V_{01|01,10}}{\partial G} = \frac{\sigma_{01,10}}{1-\sigma_{01,10}} P_{10|01,10} V_{01|01,10} > 0 \tag{10}$$

Combined, these two effects push consumers who value grocery stores less away from choosing trips that include that grocery store, either alone or combined with coffee, in favor of trips only to the coffee shop.

In total, whether the consumer goes relatively more to the coffee shop alone over the neither trip when the grocery store value falls depends on the combined effects of differential changes to the pair weighted utilities for the coffee alone trip and the neither trip. The full effects of the fall in grocery store value are summarized as

$$\begin{aligned} -\frac{\partial \ln\left(\frac{P_{01}}{P_{00}}\right)}{\partial G} = & \frac{\frac{\sigma_{01,11}}{1-\sigma_{01,11}} P_{11|01,11} V_{01|01,11} + \frac{\sigma_{01,10}}{1-\sigma_{01,10}} P_{10|01,10} V_{01|01,10}}{\sum_{gc' \neq 01} V_{01|01,gc'}} \\ & - \frac{\frac{\sigma_{00,11}}{1-\sigma_{00,11}} P_{11|00,11} V_{00|00,11} + \frac{\sigma_{00,10}}{1-\sigma_{00,10}} P_{10|00,10} V_{00|00,10}}{\sum_{gc' \neq 00} V_{00|00,gc'}}. \end{aligned} \tag{11}$$

To see the specific impact of trip-chaining, further examine the expression in 9. Unsurprisingly, consumers substitute more to the coffee alone trip where those trips are more valuable (high $V_{01|01,11}$). But, they also substitute more where consumers frequently choose the chained trip (high $P_{11|01,11}$), because a fall in grocery store value breaks more trip chains. Furthermore, the extent of these effects depends on the relative benefits of trip-chaining. For example, the change in the weighted trip value for a fall in coffee shop distance,

$$-\frac{\partial V_{01|01,11}}{\partial d^c} = \frac{\tau}{1-\sigma_{01,11}} \Big[\frac{1}{d^c} - \sigma_{01,11} \Big[\frac{1}{d^c} + \Big[\frac{1}{d^g + d^c + d^b} - \frac{1}{d^c}\Big] P_{11|01,11}\Big] V_{01|01,11} > 0. \tag{12}$$

Thus the coffee shop alone trip is not only more valuable when the coffee shop is nearby, but also to the extent that proximity increases the marginal gap between the chained and coffee alone trip distances, $\frac{1}{d^g+d^c+d^b} - \frac{1}{d^c}$.[35,36]

These results are in contrast to the log of the relative probability under the classic logit assumption,

$$\ln\left(\frac{P_{01}}{P_{00}}\right) = C + \tau \ln(2d^c/0.1), \tag{13}$$

which forces the substitution between the two trips to remain constant when grocery store values fall with respect to the features of the grocery store alone and chained trips. The nested logit model can only partially resolve these issues. Among a four trip choice set, it would be unclear ex ante which two nests and similarity parameters would best fit the data. Furthermore, any choice would necessarily restrict possibly important features of differential inter-nest substitution.[37]

I treat the adoption of online grocery platforms as a shock that causes a discrete change in the value of the grocery store for the consumer from G to G'. Logically, consumers who become online grocery shoppers have less need to visit a grocery store in person now that they have a ready supply of groceries at home.

4.C Online Choice

I separately model the availability of an online grocery platform as equivalent to gaining access to the top half of the decision tree in Figure 11. The consumer is now able to pick being an online grocery shopper and what offline trips she would make as an online grocery shopper. The choice depends on the value the consumer gets from the online grocery platform and her expected value of the daily trips she will make conditional on that choice. I set the consumer's utility maximization problem for online grocery platform adoption as

$$\max_{O_i} U_i = G^p \mathbb{1}_{\{O_i=1\}} + I_{G,C}(O_i) + \mu_i(O_i), \tag{14}$$

where U_i is the long-term utility of the consumer, G^p is the value of the online grocery platform, $I_{G,C}(O_i)$ is the inclusive value of the set of offline trip pairs as a function of online platform use, and $\mu_i(O) \sim$ EV type 1 is the taste shock for being an online grocery shopper.[38] Intuitively, platform adopters are those consumers who have a higher value for an online grocery platform and the trip

[35]Note, however, that the overall effect on the expression in 9 of a fall in coffee shop distance is ambiguous. This is because where coffee shops are closer, consumers are also relatively less likely to choose the chained trip from the coffee alone and trained trip pair in the first place, meaning $-\frac{\partial P_{11|01,11}}{\partial d^c} < 0$.

[36]Appendix C shows how an additional coffee shop changes equation C7 to, for example, allow breaking a chained trip with one coffee shop in favor of a trip to the second coffee shop alone.

[37]Appendix C provides additional details on the patterns captured in a logit and nested logit specification.

[38]During my sample period, curbside pickup of groceries purchased online was not widespread, so there are no travel costs associated with platform adoption. The recent expansion of such offerings would provide only in-store time-savings and not travel time-savings.

pairs they would choose as an online grocery shopper as compared to the value of the trip pairs they would choose otherwise. Formally, consumers who adopt an online grocery platform have

$$G^p + \mu_i(O_i) > I_{G,C}(O_i = 0) - I_{G,C}(O_i = 1) = \ln\Big[\frac{\sum_{gc}\sum_{gc'\neq gc} V_{gc}(O_i = 0)}{\sum_{gc}\sum_{gc'\neq gc} V_{gc}(O_i = 1)}\Big]. \tag{15}$$

This closes out the details of the model. This version with two stores and a binary online grocery choice is minimally sufficient for capturing the salient features of consumer substitution I observe in the data. Richer versions of the model could be developed to capture more of the mechanisms through which online retail can affect consumer choices. These could include consumer spending decisions and adjustment on the supply side. Such a model would require stronger assumptions because of card data limitations (e.g. unobserved prices and quantities). Furthermore, separate identification of the role of each mechanism would more heavily rely on the structure of the model over empirical identification. The benefit, however, would be to capture the fuller impact of online retail for consumers and producers. I leave this to future research.

4.D Parameter Identification

To examine the variation which mechanically identifies the model parameters, I lay out a progressively detailed example with consumer and retail heterogeneity. First, imagine a world without online groceries in which consumers are identical in location and income. Also assume that both the stores and the consumers are at the same location, implying that all trip costs are zero. In this world, three trip probabilities provide useful information: (1) the probability of going only to the grocery store, (2) the probability of going only to the coffee shop, and (3) the probability of going to both. The first two probabilities recover the mean values for each of the stores, G and C. Combining this information with the third probability determines the fixed benefit (or cost) of that combination.

Next, we create variation in consumer trip costs to identify the opportunity cost of distance and trip substitutability parameters. To do so, randomly locate the stores and consumers across space. Each home location generates unique trip costs specific to the consumer because no two locations provide the same set of distances between the consumer and stores. The opportunity cost of time, τ, can now be identified by the rate at which consumers decrease the probability of choosing high-cost trips in favor of low-cost trips.[39] Furthermore, the degree to which one trip is substituted for another at different distance costs identifies the substitutability parameter for any two trips.

Finally, randomly separate consumers into high- and low-income groups. Separate values for the stores, the fixed benefit of the chained trip, and the opportunity cost of time are identified

[39]Note that τ cannot be identified only from comparing consumers at, for example, different distances to the coffee shop, because consumers at different coffee shop distances will simultaneously differ in their distances to the grocery store. Identification of τ, therefore, depends on controlling for all trip costs.

separately with the data from low- and high-income customers.

The variation discussed to this point is sufficient to identify the parameters of the model when all consumers value the grocery store at an initial G. However, consumer and store locations and consumer income are not random, such that the identified parameters would be the outcomes of an endogenous process rather than the true values determining trip choices. Therefore, to help identify the model parameters, I leverage the same identification strategy deployed in Section 3 by estimating the model using the daily trip choices of early adopters and their matched controls. After an exogenous drop in grocery store value from G to G' for early adopters, we can observe how their trip choices change in response. In particular, observing the rate of substitution between two trips holding distance costs fixed is a rich source of variation for the identification of the opportunity cost of time and trip substitutability parameters separate from the store values.

To capture this variation, I include post adoption and early adoption indicators, and each interacted with an indicator for a high-income consumer in the trip utility function,

$$\begin{aligned} V_{it}(g,c) =& \beta_0 + \tau_l \ln(d_i(g,c)) \times (1-HI_i) + \tau_h \ln(d_i(g,c)) \times HI_i + \beta_1 Post_t + \\ & \beta_2 EA_i + \beta_3 HI_i + \beta_4 Post_t \times EA_i + \beta_5 Post_t \times HI_i + \beta_6 EA_i \times HI_i + \\ & \beta_7 Post_t \times EA_i \times HI_i, \end{aligned} \tag{16}$$

where HI_i is a indicator that customer i is high-income. The $Post_t$, EA_i, $Post_t \times HI_i$, and $EA_i \times HI_i$ terms hold fixed differences across periods and consumer groups to measure how adoption of an online grocery platform affects the trip utility of low- and high-income consumers relative to their matched controls. These effects are captured by the remaining coefficients, which are mapped to the trip utility parameters, $\boldsymbol{\theta_l} = (G_l, G'_l, C_l, b_l, \tau_l)$ and $\boldsymbol{\theta_h} = (G_h, G'_h, C_h, b_h, \tau_h)$, for low- and high-income consumers, respectively.[40] Estimation also recovers five trip-pair similarity parameters, $\boldsymbol{\sigma}$ (the sixth is normalized to 0 without loss of generality).

I estimate the log of equation 7 with the trip utility defined by equation 16 via full-information maximum likelihood. In this setting it is inappropriate to use an approach that estimates parameters using model-implied market share equations via GMM, as in BLP-type models, due to the infrequent consumption of some choices. For example, in more than half of consumer-month observations, there are zero shares for the bundled choice to visit the grocery store and coffee shop. As has been discussed in both the industrial organization and urban literatures, current best practices for such settings include estimation via maximum likelihood, using moment inequalities for shares, or modeling the source of zero market shares directly (Quan and Williams 2018; Dingel and Tintelnot 2020; Dubé et al. 2021; Gandhi et al. 2022).

Estimation uses the daily trip choices for a subset of the early platform adopters and their matched controls. I restrict the sample to those consumers who choose each of the four trip types at least twice in the year prior to adoption and whose median distance traveled on those trip types

[40] For example, for the grocery alone trip, β_0 identifies G_l, $\beta_0 + \beta_3$ identifies G_h, and $\beta_0 + \beta_4$ identifies G'_l

is less than 50 miles.[41] I then use the median distance traveled by each consumer to approximate the distance costs for each trip.[42]

To conclude, I note that in a model with bundled options, consumers are implicitly maximizing the choice of other stores and products not included in the model alongside their choice of grocery store and coffee shop. Relative trip utilities, therefore, depend on interactions captured by the model and those that come through un-modeled, joint consumption decisions. Therefore, estimates are conditional on the set of alternative offline and online goods available in the market. Similarly, other dimensions of choice, such as the amount consumed, are not modeled and assumed fixed conditional on market conditions.[43]

4.E Estimates

Trip utility: The parameters mapped from the MLE estimation of the PCL model are in Table 1.[44] First, note that the estimates of grocery store and coffee shop value, G and C, are negative with $C < G$ for both low- and high-income consumers. This makes sense when the value of the neither trip is the most frequent trip with a value normalized to 0. For low-income consumers each trip type is less frequent, which translates to lower store values for that group. Second, the grocery store value for high-income consumers who use online grocery platform, G'_h, is 7.4% lower than G_h. Low-income consumer grocery store values fall slightly farther, by 7.6%. These declines closely match the decline in trip frequency from Figure 3. Third, the combined trip fixed benefit, b, is positive for both consumer types, reflecting that visiting both stores on the same trip lowers the combined trip cost relative to visiting each store on separate trips. The benefit is slightly higher for low-income consumers. Fourth, trip utility declines for both low- and high-income consumers as trip distance increases, though the magnitudes are small. This is because the opportunity costs here reflect the distance trade-off within consumers' preferred trip options rather than among all possible trips.[45]

Estimates for one nested logit model and the logit model are shown for comparison. The nested logit version shown here uses nests in which trips with coffee are in one nest and trips without are in the other. In the trip utility parameters, pre- and post-adoption grocery store values for low- and high-income consumers in alternative models are within 10% of the PCL values. From there, important differences emerge. Coffee shop values and the fixed benefits to trip chaining are far lower in the alternative models with generally stronger disutility to distance for each type of consumer. These differences are driven by the similarity parameters. The PCL model allows for variation in

[41] Large distances occur when consumers live far from their reported zip code or spend substantial time away.

[42] To avoid taking logs of zero distances, on trips when consumers or stores are in the same zip code I assume that within zip code distances are equal to half the radius of a circle of the same area as the zip code.

[43] Griffith et al. (2009) and Thomassen et al. (2017) study non-travel related choice dimensions in the context of grocery store purchases.

[44] See Table C4 for the estimation results from equation 16 used in the mapping.

[45] For example, in their study of consumption across grocery stores, Thomassen et al. (2017) find a coefficient on grocery store distance of 0.4.

Table 1: PCL and Logit Model Parameters

Trip utility:	*PCL*	*NL1*	*Logit*	***Similarity:***	*PCL*	*NL1*	*Logit*
G_h	-0.760	-0.679	-0.806	$\sigma_{01,10}$	0.974		0
G_l	-0.994	-0.823	-1.038	$\sigma_{01,00}$	0.910		0
				$\sigma_{01,11}$	0.565	0.505	0
G'_h	-0.816	-0.755	-0.908	$\sigma_{10,11}$	0.733		0
G'_l	-1.101	-0.942	-1.207	$\sigma_{00,11}$	0.757		0
				$\sigma_{10,00}$	0	0.194	0
C_h	-0.816	-1.844	-2.037				
C_l	-1.013	-1.772	-1.918				
b_h	0.330	0.324	0.106				
b_l	0.520	0.341	0.024				
τ_h	-0.010	-0.005	-0.014				
τ_l	-0.017	-0.034	-0.048				

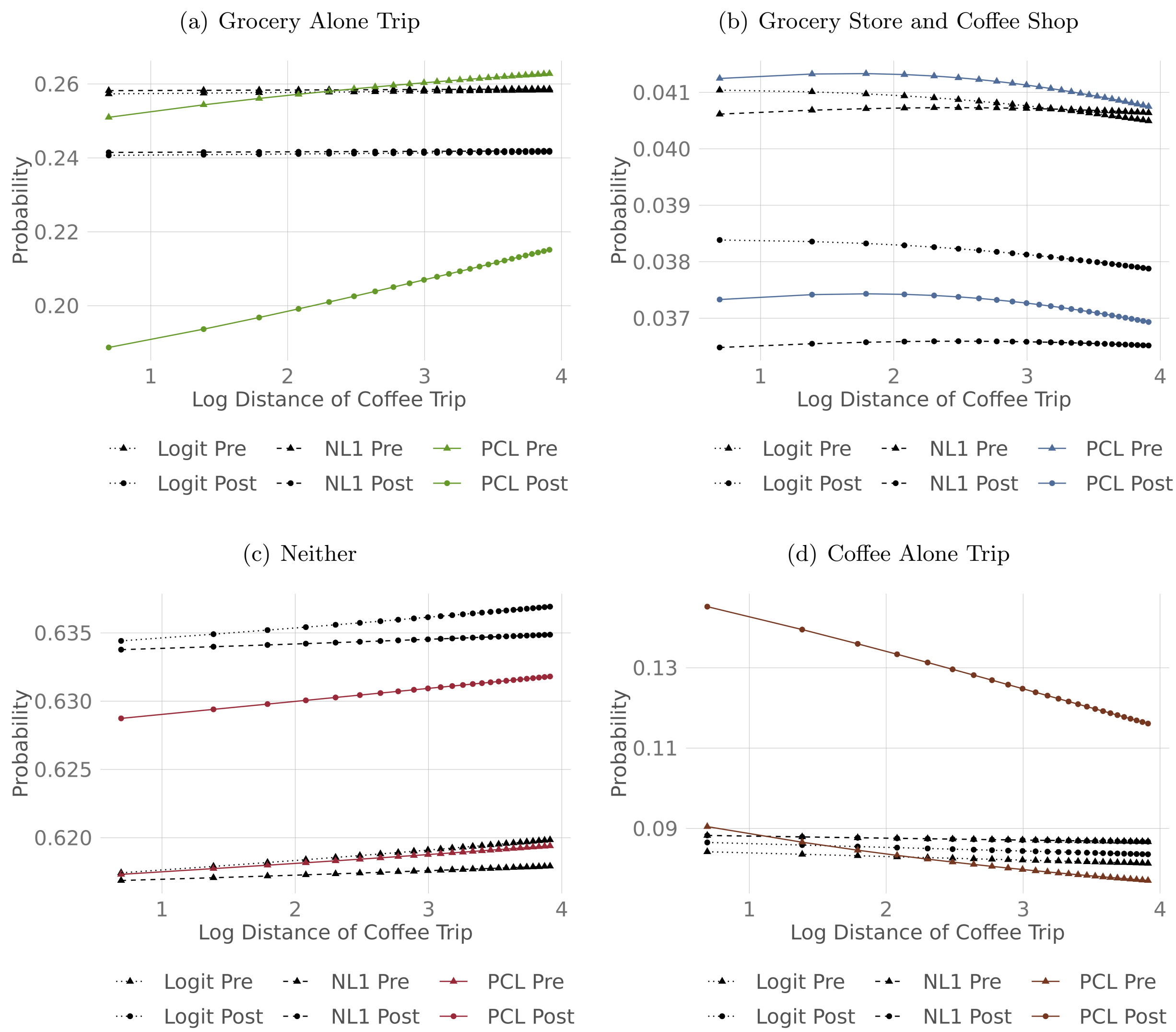
Notes: This table shows the parameters mapped from MLE estimates for the PCL, one possible nested logit model (NL1), and the logit model. The nested logit model shown here nests the two trips including the coffee shop in one nest and the two without in the other. See Table C4 for estimates and details on mapping to the parameters.
Source: To be in the model estimation sample, the consumer (1) must make 2 of each trip type in the year before adoption and (2) have median distance costs for each trip type of less than 50 miles during that year. The first requirement focuses the estimation on consumers who trade-off utility from all four trip bundles in their daily trip decision. The latter requirement eliminates consumers who do not regularly live at their reported home zip code. There are 8,604 early adopters and matched controls who meet these requirements.

the strength of substitution patterns across different pairs of trips driven by preferences rather than primarily relying on trip features to explain trip choice. The results show that the grocery store alone and coffee shop alone trips are highly similar, reflecting consumers' tendency to substitute primarily from the former to the latter in the reduced form estimates. But, fixed arbitrarily at 0 for the logit and unmodelled in the nested logit, trip utility parameters must adjust to rationalize observed choices.

Simulation: In a short exercise, I demonstrate the importance of estimating the similarity parameters. This highlights that models using independent shocks in discrete choices can appear to perform well in the cross-section, but may miss important substitution patterns when model parameters change, such as in counterfactual exercises. The implication is that a wide class of modern spatial equilibrium models and CES consumption models may assume independent extreme value distributional shocks for convenience at some real cost.

I generate a set of hypothetical high-income consumers who are equidistant to the grocery store but vary in their distance to the coffee shop. This generates customers with the same distance costs for the grocery alone trip and varying distance costs for the coffee alone and combined trips. Figure 12 shows the resulting variation in the PCL, nested logit, and logit trip probabilities before and after platform adoption by customer distance to the coffee shop. In general, we see that because the opportunity costs to trips are small, slopes are fairly flat and differences across distance are driven more by the strength of the trip similarity parameters. For the trip probabilities before adoption, there are some subtle differences across model specifications. For instance, compared to the logit model, the nested logit model (in which the two trips which include coffee are nested together) predicts slightly more substitution toward the chained trip over the trip to neither store when the coffee shop is very far away. This stronger substitution is driven by the nesting of the two trips which include the coffee shop. However, the PCL model is better able to capture that both trips should be much less likely when the coffee shop is far away. Despite this, one might feel that any model is reasonable given the roughly similar probabilities for each trip prior to platform adoption.

Figure 12: Substitution Patterns Generated by the PCL and Alternative Models



Notes: This figure shows the varying substitution patterns generated by the PCL and alternative models for the grocery alone and coffee alone trips using a set of simulated high-income customers who vary only in their distance to the coffee shop. The nested logit model nests the two trips with coffee in one nest and the two trips without coffee in the other. While all three models give similar probabilities in the cross-section prior to adoption, only the PCL can generate substitution patterns that vary with coffee shop distance when consumers' value for the grocery store falls post-adoption.

However, the models show wide differences in the predicted substitution patterns after adoption. Only the PCL model matches the intuitive changes in the distance gradients observed in the reduced form results in Figure 10. This is because, in the logit model, the neither trip and coffee alone trip probabilities must shift in constant proportion as they are irrelevant alternatives when the grocery store value changes. The nested logit model does little to resolve these issues, as when irrelevant alternatives are maintained across nests, this continues to impose heavy restrictions on substitution patterns. In this nested logit model, for example, the elasticity of the relative share of the coffee alone trip to the chained trip with respect to a fall in grocery store value must be constant.[46] But by allowing substitutions between different trips to dominate more or less at difference coffee shop distances, the PCL model can predict that consumers will substitute more to the coffee alone trip when grocery store values fall, and that substitution will be disproportionately higher when they are closer to the coffee shop.

Online platform values: The final parameter to be estimated in the model is the value of the online grocery platform. To do this, note that the total welfare gain from the platform, the change in compensating variation, equates to the log sum gain in value from the platform.

$$\Delta CV_i = \ln\Big[\frac{exp(G^p + I_{G,C}(O_i = 1)) + exp(I_{G,C}(O_i = 0))}{exp(I_{G,C}(O_i = 0))}\Big], \tag{17}$$

which reduces to a simple function of the probability of online grocery platform adoption,

$$\Delta CV_i = \ln\Big[\frac{1}{1 - P_{O_i}}\Big], \tag{18}$$

where P_{O_i} is the probability of adoption of the online grocery platform. Intuitively, the higher the probability of adoption, the higher the increase in welfare to the consumer from being able to be an online grocery shopper.[47] I replace the probability of adoption with the share of consumers that adopt a platform in equation 18 and combine with equation 17 to solve for the value of the online grocery platform in each zip code,

$$G^p = I_{G,C}(O_i = 0) - I_{G,C}(O_i = 1) + \ln\Big[\frac{s_{O=1}}{1 - s_{O=1}}\Big], \tag{19}$$

where $s_{O=1}$ is the share adopting in a zip code. We see that platform values rationalize the leftover variation in adoption rates observed in the data after the variation in differences in offline trip values conditional on platform adoption.

The ingredients needed to solve for this value are the inclusive value of offline trips conditional on platform adoption and adoption rates. I calculate the inclusive values for offline trips using

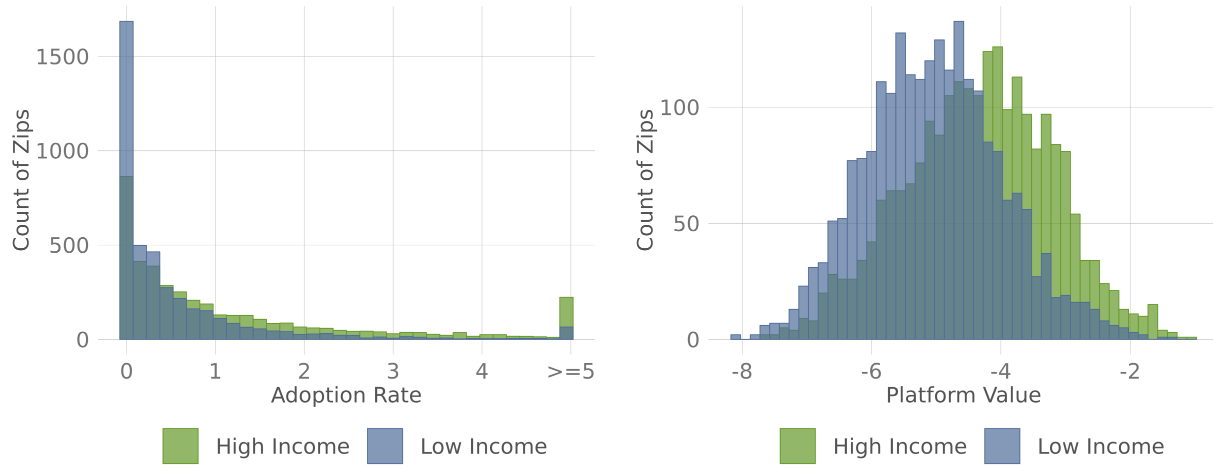
[46] This is not a feature of the particular nested logit model in Figure 12, but of the restrictions created with any IIA assumption in this setting. See Figure C1 for a comparison of the three possible nested logit structures.

[47] This formulation falls out of the simple structure of the first stage of the discrete choice, and is not specific to the PCL structure in the second stage.

Figure 13: Grocery Platform Adoption and Value

(a) Adoption Rates by Income

(b) Values by Income



Notes: Panel (a) shows the distribution of adoption rates by low- and high-income consumers across zip codes. A small share of customers adopt a platform during the sample window in most zip codes. Panel (b) shows the values for the platforms implied by the adoption rates using equation 19.
Source: Author's calculations using zip codes with at least 500 customers from the panel. There are 4,159 zip codes in total. Platform values in Panel (b) are only derived for the 2,276 zip codes where adoption rates for low- and high-income consumers are both positive. Equation 19 cannot reconcile zero adoption rates, with the result that the left tails of the distributions in Panel (b) are limited.

estimated parameters from the PCL model and the median distance traveled on each of the four trip types by all consumers in a zip code.[48] Figure 13 Panel (a) shows adoption rates by income group. Even though the aggregate adoption rate in the population is small, there are many zip codes in which a much higher share of customers adopt a grocery platform, particularly among high-income consumers. Figure 13 Panel (b) shows the resulting distribution of online grocery platform values for each income group where adoption rates are above zero. They are, on average, negative, reflecting that non-adopters outweigh adopters in each case. However, the platform value distribution for the high-income consumers is shifted substantially to the right, driven by their higher adoption rates for platforms.

5 Welfare and Counterfactuals

5.A Consumer Welfare Gains

I use the correspondence between welfare gains and adoption rates to look at important sources of variation in welfare gains across space. I find that welfare gains are highest for the consumers who would benefit the most from the time saved making offline trips, as characterized by the model. The

[48] Figure C3 shows the distribution of offline trip values for low- and high-income consumers for those that do and do not adopt an online grocery platform across zip codes.

first column of Table 2 shows that welfare gains strongly increase with zip code median income. Zip codes with median incomes in the fifth quintile nationally experience welfare gains three times those in the first quintile (results not shown in table). Controlling for income, measures of grocery and coffee access in a zip code are also correlated with platform adoption and welfare gains. Counter to the model, welfare gains are higher in places with more grocery density and shorter grocery store alone trips, but this is driven by the non-random location of stores (columns (2) and (3)) in dense urban areas where incomes and adoption rates are also high. In column (4), controlling for both income and zip code store density, zip codes with longer grocery store alone and coffee store alone trips have higher adoption and welfare gains. Interestingly, zip codes where consumers are willing to travel farther for chained trips to both the grocery store and coffee shop have lower adoption rates and welfare gains. This gives further support to the theme that valuable chained trips can insulate offline retail from online retail competition.

Table 2: Welfare Gains by Zip Code Features

	(1)	(2)	(3)	(4)
Intercept	-0.1870	-0.1874	-0.1404	-0.1672
	(0.0077)	(0.0075)	(0.0078)	(0.0078)
Median Income	0.0177	0.0182	0.0160	0.0177
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Grocery Density		0.0031		0.0027
		(0.0002)		(0.0002)
Coffee Density		0.0015		0.0012
		(0.0002)		(0.0002)
Median Grocery Alone Trip Distance			-0.0017	0.0011
			(0.0005)	(0.0005)
Median Coffee Alone Trip Distance			0.0043	0.0039
			(0.0007)	(0.0006)
Median Both Trip Distance			-0.0145	-0.0110
			(0.0010)	(0.0010)
Observations	4159	4159	4159	4159
Adjusted R^2	0.1370	0.2639	0.2139	0.2867

Notes: Standard errors in parentheses. This table shows the correlations between zip code welfare from online grocery platform availability and zip code median income and retail features. Covariates are measured in logs. Store density is measured as the number of grocery stores and coffee shops operating per square mile at the start of the panel. Trip costs are measured as the median distance traveled by any customer in the zip code for that trip at the start of the panel. Average platform welfare is 0.011.
Source: Author's calculations using the adoption rates of customers living in 4,159 zip codes with at least 500 customers. Zip code median income is from the 2014-2018 American Community Survey.

5.B Counterfactuals

I use the model to quantify the impact of a larger online retail market on offline brick-and-mortar stores and develop strategies that offline competitors can use to compete against its growth. Crucially, these exercises show that there are trade-offs in the competitive location strategies of brick-and-mortar stores. In the model, grocery stores can benefit from being close to coffee shops through chained store visits; but neighborhoods with nearby coffee shops provide the most valuable non-grocery trips, making platform adoption more likely in those places. In considering where to locate additional stores, therefore, offline retailers must balance the positive spillover effects of density with other stores against the tendency of consumers in places with valuable local alternative trip options to more strongly prefer online shopping.[49]

Larger online grocery market: A much larger online grocery market in the coming years is a near certainty. The low adoption rates during my sample period likely reflect slow adoption typical in new product markets and early platform quality issues, rather than a large-scale rejection of platforms by most consumers. The recent pandemic will further accelerate the arrival of the more mature market.[50] To create this counterfactual future, I start by increasing the platform value for each zip code and income group by 50%. Figure 14 shows the counterfactual distributions of adoption rates under this scenario. At the mean values, 4.4% of low-income consumers and 7.5% of high-income consumers adopt an online grocery platform in a zip code. Particularly for high-income consumers, adoption rates in excess of 20% in a zip code are not uncommon.

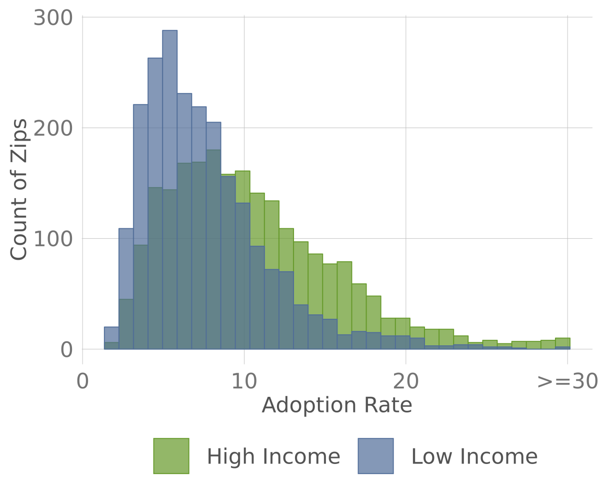
In such an environment with higher adoption rates, I forecast changes in trip-frequency for grocery stores and coffee shops for the population in each zip code (including platform adopters and non-adopters). Figure 15 shows the distribution in forecasted trip changes by income group. For both store types, there is wide dispersion in trip frequency changes for low- and high-income consumers. The distributions of declines for grocery stores are relatively similar for both groups, as they experience similar declines in their grocery store values. However, high-income consumers substitute more strongly toward coffee shops, given their higher value for coffee shops. Total zip code effects from higher platform values are derived from the effects for low- versus high-income consumers weighted by income group populations in each zip code. The results are in the first three rows of the second panel of Table 3. With 50% higher platform rates, mean adoption rates jump 7.6 percentage points with a 1.8% decline in grocery store trips and a 2.8% increase in coffee shop trips.

Given this counterfactual future, I use the model to quantify the effectiveness of strategies that offline competitors may use to make trips to their stores more attractive and reduce increased adoption of online alternatives. These exercises show that, at least in the grocery market where

[49] Of course, where grocery stores and online grocery platforms are owned by the same firm, cross-channel spillovers mean that some decline in in-store sales may be preferable if overall purchases rise, either from online grocery shoppers buying more groceries overall or expanding their customer base.

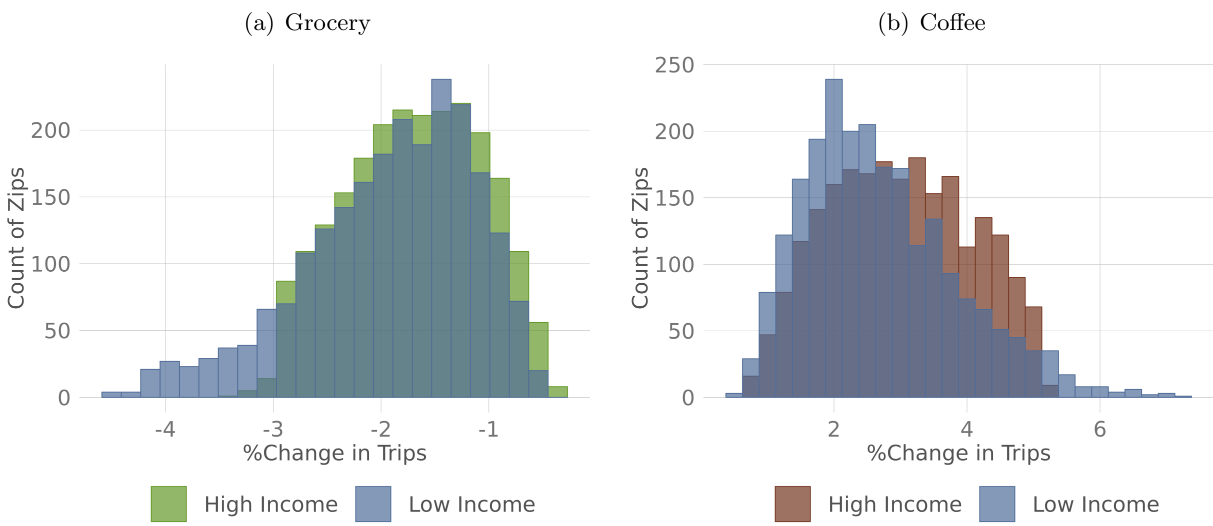
[50] Relihan et al. (2020) shows that online grocery spending in the early months of the pandemic more than doubled.

Figure 14: Predicted Platform Adoption Rates



Notes: This figure shows the predicted rate of platform adoption for low- and high-income groups across zip codes with 50% higher platform values than those in Figure 13 Panel (b).
Source: Author's calculations using zip codes with at least 500 customers from the panel and positive adoption rates for both low- and high-income consumers. There are 2,276 such zip codes.

Figure 15: Predicted Trip Changes with Higher Platform Values



Notes: This figure shows the distribution of predicted percent changes in grocery store and coffee shop trip frequencies across zip codes given 50% higher online grocery platform values than those in Figure 13 Panel (b).
Source: Author's calculations using zip codes with at least 500 customers from the panel and positive adoption rates for both low- and high-income consumers. There are 2,276 such zip codes.

Table 3: Counterfactuals

Base	*Outcome*	*Statistic*					
		Mean	St. Dev.	Min	P25	P75	Max
	Adoption Rate	1.66	2.15	0.04	0.43	2.05	23.27
	Grocery Trips	8.32	0.38	7.11	8.06	8.61	9.26
	Coffee Trips	3.75	0.14	3.13	3.68	3.80	4.44
Counterfactual	*Outcome*	*Statistic*					
		Mean	St. Dev.	Min	P25	P75	Max
$G^p \uparrow 50\%$							
	Adoption Rate Δ	7.63	2.88	1.86	5.33	9.78	14.11
	Grocery Trips $\%\Delta$	-1.75	0.63	-3.27	-2.20	-1.25	-0.45
	Coffee Trips $\%\Delta$	2.84	1.03	0.71	2.03	3.63	5.26
$+\ b \uparrow 10\%$							
	Adoption Rate Δ	7.62	2.88	1.86	5.32	9.77	14.09
	Grocery Trips $\%\Delta$	-0.94	0.66	-2.45	-1.43	-0.42	0.44
	Coffee Trips $\%\Delta$	4.99	0.98	2.93	4.23	5.75	7.57
$+\ G\ \&\ G' \uparrow 5\%$							
	Adoption Rate Δ	7.56	2.86	1.84	5.28	9.71	14.00
	Grocery Trips $\%\Delta$	6.65	1.54	3.52	5.50	7.56	14.26
	Coffee Trips $\%\Delta$	-7.47	2.30	-17.07	-8.94	-5.81	-1.87
$+\ d_i(1,0) \downarrow 50\%$							
	Adoption Rate Δ	7.60	2.87	1.86	5.31	9.76	14.09
	Grocery Trips $\%\Delta$	0.65	1.41	-4.40	-0.36	1.59	5.37
	Coffee Trips $\%\Delta$	-1.40	2.41	-8.35	-3.04	0.31	7.52

Notes: The top panel in this table table shows summary statistics for actual adoption rates and predicted grocery store and coffee shop trip frequencies given model parameters and median trip distances traveled by all customers in a zip code. Trip frequencies are measured as days per month. The second panel shows the predicted changes in platform adoption rates and percent changes in grocery store and coffee shop trip frequencies across zip codes under four counterfactuals. See text for details.
Source: Author's calculations using model parameters and zip codes with at least 500 customers from the panel and positive adoption rates for both low- and high-income consumers. There are 2,276 such zip codes.

offline trips are only partially replaced by online retail, marginal changes that affect consumer trip choice can blunt or reverse the impact of online retail competition. The remainder of the second panel in Table 3 shows the independent effect of each.

Increasing the fixed benefit to chained trips: One strategy that grocery stores can pursue to compete with online retail is to tie themselves more closely to services. These are less substitutable with online alternatives than goods and, as the reduced form results show, more likely to be purchased with the time saved from online retail. In terms of the model, one way to do this is to increase the fixed benefit to the chained trip, b. This would partly reflect a strategy in which grocery stores open an internal, but still independent coffee shop, removing both distance and other travel inconveniences from two stops. A modest 10% increase in the fixed benefits for both low- and high-income consumers has wide benefits for both grocery stores and coffee shops. Under this counterfactual, a quarter of zip codes have grocery stores with increasing trip frequency, rather than declines. The spillover effect to coffee shops is such that their increase in trips is double that from the wider adoption of online groceries alone.

Higher grocery store values: Another strategy that a competitor to online retail can pursue is to increase their quality, such that consumers desire more frequent trips to a store. For example, instead of supporting an independent coffee shop, a grocery store could operate one itself internally. Table 3 shows the results of such a quality improvement that increases grocery store values 5%. In this scenario, additional trips to the grocery store displace trips to external coffee shops.

Lower distances to preferred grocery stores: Finally, offline competitors to online retail could improve their physical accessibility. The model used in this paper, with only one grocery store and one coffee shop, cannot quantify the impact of higher market access to consumers, but rather the effect of shorter distances to consumers for whom the store is already the preferred option. Because the between-trip distance opportunity costs are small, large changes in distances are needed to achieve meaningful changes in outcomes. To that end, I decrease the distance traveled by consumers on their grocery store alone trip by 50%, holding other distances fixed. Locating closer to consumers reverses much of the decline in trips for grocery stores from the wider adoption of online grocery platforms, but also at the expense of coffee shops. Such an exercise emphasizes that marginal improvements in accessibility alone will not generate a meaningful increase in trips for a product, like groceries, which consumers visit at frequencies determined largely by other factors.

6 Conclusion

The continuing rise of online retail will transform local offline economies and the way consumers and retailers interact. The effects of the pandemic are likely to accelerate this transformation through the rapid adoption of new online products and premature closures of many brick-and-mortar stores. The results of this paper shed light on the "new normal" that may emerge post pandemic. In that future when online retail is more dominant, not all brick-and-mortar stores are doomed. Time use substitution is a mechanism that can create offline shopping complements as well as substitutes to online retail.

The research presented here shows that the benefits from online retail to consumers will be uneven. Consumers who can afford to access online products and have high opportunity costs of time will substantially benefit from the entry of new online products. Those who live in neighborhoods with less access to goods, but with more access to services, stand to benefit further through the greater benefits in time-savings from online goods and easy access to their neighborhoods' amenities.

For firms, this research shows that offline retailers that compete directly with online retailers on product are negatively impacted, particularly those that are most costly for consumers to reach. However, those retailers can adopt strategies to make trips to their stores more attractive. These include locating more closely with time-intensive, non-tradable services, offering more services themselves, and locating more closely to consumers. However, such strategies are likely to be less effective in retail dense urban areas, where the adoption of online retail is more attractive because of the high-value of trips to time-intensive and non-tradable services. Therefore, retail firms who directly compete with online, should carefully consider the trade-offs in their location decisions. In contrast, time-intensive and non-tradable services can substantially benefit from the increase in available time that comes through the rise of online retail. Those benefits will increase with the density of similar firms, likely spurring an increase in urban consumption amenities based on services.

The implications for local offline economies go beyond which consumers buy what products where to the functioning of other sectors. These include local labor markets, in which 14.5 million people were employed in brick-and-mortar retail jobs in October 2021.[51] In addition, changing spatial consumption patterns will differentially affect the value of commercial property (Rosenthal et al. Forthcoming). The results presented here imply higher values closer to consumers, including residential neighborhoods and locations with high foot traffic, and higher values in locations with many services. Local governments may also struggle to meet their funding needs if the revenue from traditional sales taxes on goods decline more than those on services rise. While these changes may be painful, local offline economies that can transform to coexist and complement online retail will ultimately be able to improve the welfare of residents and firms.

References

Aguiar, Mark and Erik Hurst, "Life-Cycle Prices and Production," *The American Economic Review*, 2007, *97* (5), 1533–1559b.

_ and _ , "Measuring Trends in Leisure: The Allocation of Time over Five Decades," *The Quarterly Journal of Economics*, 2007, *122* (3), 969–1006.

_ , _ , and Loukas Karabarbounis, "Time Use During the Great Recession," *The American Economic Review*, 2013, *103* (5), 1664–1696.

Arentze, Theo A., Harmen Oppewal, and Harry J. P. Timmermans, "A Multipurpose Shopping Trip Model to Assess Retail Agglomeration Effects," *Journal of Marketing Research*, 2005, *42* (1), 109–115.

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

[51] This is 11.7% of private employment. Calculated using Bureau of Labor Statistics Current Employment Survey. Brick-and-mortar employment defined as retail minus non-store retail employment.

Baker, Scott R, Stephanie Johnson, and Lorenz Kueng, "Shopping for Lower Sales Tax Rates," *American Economic Association: Macroeconomics*, 2020, *Forthcoming* (Forthcoming).

Bakos, J. Yannis, "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science*, 1997, *43* (12), 1676–1692.

Baugh, Brian, Itzhak Ben-David, and Hoonsuk Park, "Can Taxes Shape an Industry? Evidence from the Implementation of the "Amazon Tax"," *The Journal of Finance*, 2018, *73* (4), 1819–1855.

Bell, David R. and Sangyoung Son, "Neighborhood Effects and Trial on the Internet: Evidence from Online Grocery Retailing," *Quantitative Marketing & Economics*, 2007, *5* (4), 361–400.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

Brandão, António, Joao Correia da Silva, and Joana Pinho, "Spatial Competition Between Shopping Centers," *Journal of Mathematical Economics*, 2014, *50*, 234–250.

Bronnenberg, Bart J., Tobias J. Klein, and Yan Xu, "Consumer Time Budgets and Grocery Shopping Behavior," Working Paper 2020.

Brooks, Charles M., Patrick J. Kaufmann, and Donald R. Lichtenstein, "Travel Configuration on Consumer Trip-chained Store Choice," *Journal of Consumer Research*, 2004, *31* (2), 241–248.

_ , _ , and _ , "Trip Chaining Behavior in Multi-destination Shopping Trips: A Field Experiment and Laboratory Replication," *Journal of Retailing*, 2008, *84* (1), 29–38.

Brynjolfsson, Erik, Yu Hu, and Mohammad S. Rahman, "Battle of the Retail Channels: How Product Selection and Geography Drive Cross-channel Competition," *Management Science*, 2009, *55* (11), 1755–1765.

Butters, Gerard R., "Equilibrium Distriubtions of Sales and Advertising Prices," *The Review of Economic Studies*, 1977, *44* (3), 465–491.

Callaway, Brantly and Pehdra H.C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225*, 200–230.

Cohen, Michael Andrew and Marc Rysman, "Payment Choice with Consumer Panel Data," Working Paper No. 13-6, Federal Reserve Bank of Boston 2013.

Compiani, Giovanni, "Market Counterfactuals and the Specification of Multi-Product Demand: A Nonparametric Approach," Working Paper 2021.

Conley, Timothy G. and Christopher R. Udry, "Learning about a New Technology: Pineapple in Ghana," *The American Economic Review*, 2010, *100* (1), 35–69.

Cosman, Jacob, "Industry Dynamics and the Value of Variety in Nightlife: Evidence from Chicago," Working Paper, University of British Columbia 2017.

Couture, Victor, "Valuing the Consumption Benefits of Urban Density," Working Paper, University of California, Berkeley 2016.

_ **, Benjamin Faber, Yizhen Gu, and Lizhi Liu**, "Connecting the Countryside via E-commerce: Evidence from China," *American Economic Review: Insights*, 2020, *Forthcoming.*

_ **, Cecile Gaubert, Jessie Handbury, and Erik Hurst**, "Income Growth and the Distributional Effects of Urban Spatial Sorting," *Review of Economic Studies*, 2021, *3* (1).

Crane, Leland D., Ryan A. Decker, Aaron Flaaen, Adrian Hamins-Puertolas, and Christopher Kurz, "Business Exit During the COVID-19 Pandemic: Non-Traditional Measures in Historical Context," Finance and Economics Discussion Series 2020-089rl, Board of Governors of the Federal Reserve System 2021.

Davis, Donald R, Jonathan I. Dingel, Joan Monras, and Eduardo Morales, "How Segregated is Urban Consumption?," *Journal of Political Economy*, 2019, *127* (4), 1684–1738.

Dellaert, Benedict G.C., Theo A. Arentze, Michel Bierlaire, Aloys W. J. Borgers, and Harry J.P. Timmermans, "Investigating Consumers' Tendency to Combine Multiple Shopping Purposes and Destinations," *Journal of Marketing Research*, 1998, *35* (2), 177–188.

Dingel, Jonathan I. and Felix Tintelnot, "Spatial Economics for Granular Settings," Working Paper No.27287, National Bureau of Economic Research 2020.

Dolfen, Paul, Liran Einav, Peter J. Klenow, Benjamin Klopack, Jonathan D. Levin, Larry Levin, and Wayne Best, "Accessing the Gains from E-commerce," Working Paper No.25610, National Bureau of Economic Research 2019.

Dubé, Jean-Pierre, Ali Hortaçsu, and Joonhwi Joo, "Random-Coefficients Logit Demand Estimation with Zero-Valued Market Shares," *Marketing Science*, 2021, *40* (4), 637–660.

Einav, Liran, Dan Knoepfle, Jonathan Levin, and Neel Sundaresan, "Sales Taxes and Internet Commerce," *American Economic Review*, 2014, *104* (1), 1–26.

Ellison, Glenn and Sara Fisher Ellison, "Tax Sensitivity and Home State Preferences in Internet Purchasing," *American Economic Journal: Economic Policy*, 2009, *1* (2), 53–71.

Forman, Chris, Anindya Ghose, and Avi Goldfarb, "Competition Between Local and Electronic Markets: How the Benefit of Buying Online Depends on Where you Live," *Management Science*, 2009, *55* (1), 47–57.

Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi, "Estimating Demand for Differentiated Products with Zeroes in Market Share Data," Working Paper 2022.

Gentzkow, Matthew, "Valuing New Goods in a Model with Complementarity," *The American Economics Review*, 2007, *97* (3), 713–744.

Glaeser, Edward L., Jed Kolko, and Albert Saiz, "Consumer City," *Journal of Economic Geography*, 2001, *1* (1), 27–50.

Goolsbee, Austan, "In a World Without Borders: The Impact of Taxes on Internet Commerce," *The Quarterly Journal of Economics*, 2000, *115* (2), 561–576.

_ **and Peter J. Klenow**, "Evidence on Learning and Network Externalities in the Diffusion of Home Computers," *Journal of Law and Economics*, 2002, *45* (2), 561–576.

Gorback, Caitlin, "Ridesharing and the Redistribution of Economic Activity," Technical Report 2020.

Griffith, Rachel, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo, "Consumer Shopping Behvaior: How Much Do Consumers Save?," *Journal of Economic Perspectives*, 2009, *23* (2), 99–120.

Handbury, Jessie and David E. Weinstein, "Goods Prices and Availability in Cities," *Review of Economic Studies*, 2015, *82* (1), 258–296.

Harwitz, Mitchell, Barry Lentnek, and Subhash C. Narula, "Do I Have to go Shopping Again? A Theory of Choice with Movement Costs," *Journal of Economic Geography*, 1983, *13* (2), 165–180.

Holmes, Thomas J, "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 2011, *79* (1), 253–302.

Hotelling, Harold, "Stability in Competition," *The Economic Journal*, 1929, *39* (153), 41–57.

Houde, Jean-François, "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *The American Economic Review*, 2012, *102* (5), 2147–2182.

Jardim, Eduardo, "All in the Mix: Spillovers and the Agglomeration of Neighborhood Retail," Technical Report 2015.

Jin, Ginger Zhe and Andrew Kato, "Dividing Online and Offline: A Case Study," *The Review of Economic Studies*, 2007, *74* (3), 981–1004.

Jingting, Fan, Lixin Tang, Weiming Zhu, and Ben Zou, "The Alibaba Effect: Spatial Consumption Inequality and the Welfare Gains from E-commerce," *Journal of International Economics*, 2018, *114*, 203–220.

Jo, Yoon J., Misaki Matsumura, and David E. Weinstein, "The Impact of E-Commerce on Relative Prices and Consumer Welfare," Working Paper No.26506, National Bureau of Economic Research 2019.

Koppelman, Frank S. and Chieh-Hua Wen, "The Paired Combinatorial Logit Model: Properties, Estimation and Application," *Transportation Research Part B: Methodological*, 2000, *34* (2), 75–89.

Mehta, Nitin and Yu Ma, "A Multicategory Model of Consumers' Purchase Incidence, Quantity, and Brand Choice Decisions: Methodological Issues and Implications on Promotional Decisions," *Journal of Marketing Research*, 2012, *49* (4), 435–451.

Miyauchi, Yuhei, Kentaro Nakajima, and Stephen J. Redding, "Consumption Access and Agglomeration: Evidence from Smartphone Data," Working Paper 2022.

Narula, Subhash C., Mitchell Harwitz, and Barry Lentnek, "Where Shall we Shop Today? A Theory of Multiple-stop, Multiple-purpose Shopping Trips," *Papers of the Regional Science Association*, 1983, *53* (1), 159–173.

Nevo, Aviv and Arlene Wong, "The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession," *International Economic Review*, 2019, *60* (1), 25–51.

Pozzi, Andrea, "Shopping Cost and Brand Exploration in Online Grocery," *American Economic Journal: Microeconomics*, 2012, *4* (3), 96–120.

Quan, Thomas W. and Kevin R. Williams, "Product Variety, Across-market Demand Heterogeneity, and the Value of Online Retail," *RAND Journal of Economics*, 2018, *49* (4), 877–913.

Relihan, Lindsay E., Marvin M. Ward, Chris W. Wheat, and Diana Farrell, "The Early Impact of COVID-19 on Local Commerce: Changes in Spend Across Neighborhoods and Online," *Covid Economics*, 2020, *28*, 1–28.

Rosenthal, Stuart S., William C. Strange, and Joaquin A. Urrego, "Are City Centers Losing Their Appeal? Commercial Real Estate, Urban Spatial Structure, and COVID-19," *Journal of Urban Economics*, Forthcoming.

Serra, Daniel and Rosa Colomé, "Consumer Choice and Optimal Locations Models: Formulations and Heuristics," *Papers in Regional Studies*, 2001, *80* (4), 439–464.

Sinai, Todd and Joel Waldfogel, "Geography and the Internet: Is the Internet a Substitute or a Complement for Cities?," *Journal of Urban Economics*, 2004, *56* (1), 1–24.

Su, Yichen, "Measuring the Value of Urban Consumption Amenities: A Time-use Approach," MPRA Working Paper 113158 2022.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225*, 175–199.

Thomassen, Øyvind, Howard Smith, Stephan Seiler, and Pasquale Schiraldi, "Multi-category Competition and Market Power: A Model of Supermarket Pricing," *American Economic Review*, 2017, *107* (8).

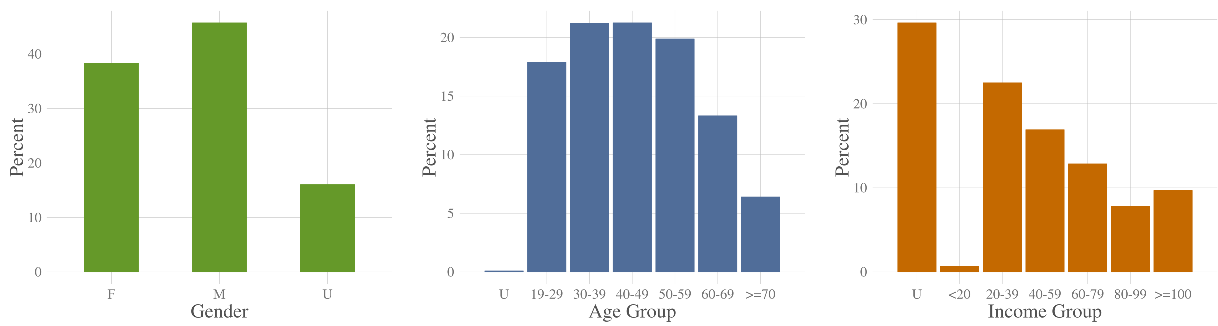
Ushchev, Philip, Igor Sloev, and Jacques-François Thisse, "Do we go Shopping Eowntown or in the Burbs?," *Journal of Urban Economics*, 2015, *85*, 1–15.

Welander, Tom, "Trends in Consumer Payments and Retail Banking: Report 1 of 4," Technical Paper, GC Insights Marketing Research Services 2014.

7 Appendix A: Summary Statistics

Figure A1: Panel Demographics

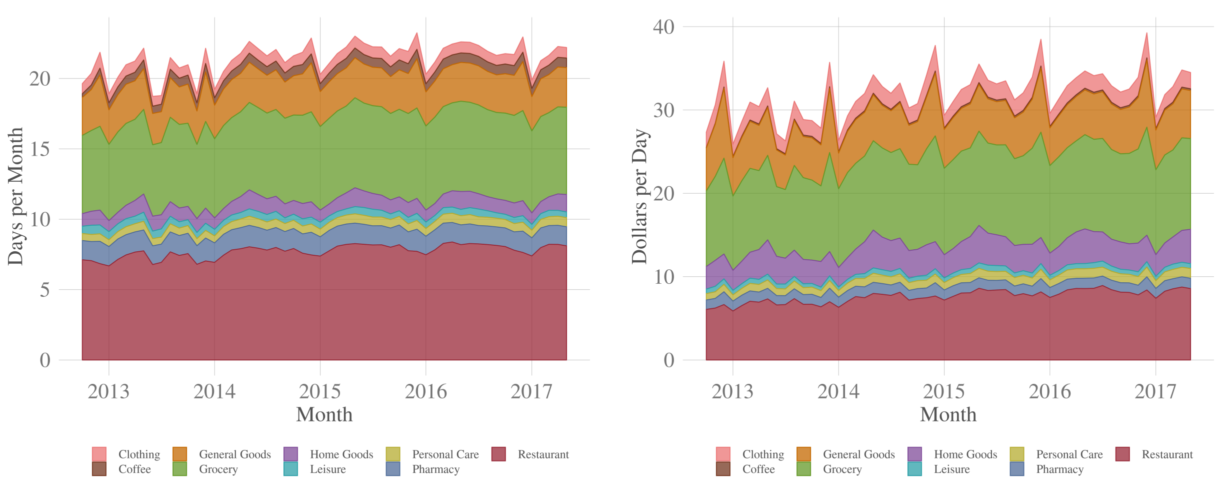
(a) Sex (b) Age (c) Income



Notes: This figure shows the gender, age, and income distributions in the panel. Uncategorized customers in a category are labeled with a "U". Sex is inferred from customer names and skews male. Customers younger than 18-years-old are excluded. Income is estimated for deposit customers in thousands of dollars using a variety of customer-reported inputs, such as income on a mortgage application. There is no estimated income for credit-only customers, about 30% the sample. For classification into low- and high-income consumers, these credit-only customers are treated as high-income.
Source: Author's calculations using the card transactions from the base 7.7 million customer sample.

Figure A2: Trips and Spending for Offline Products

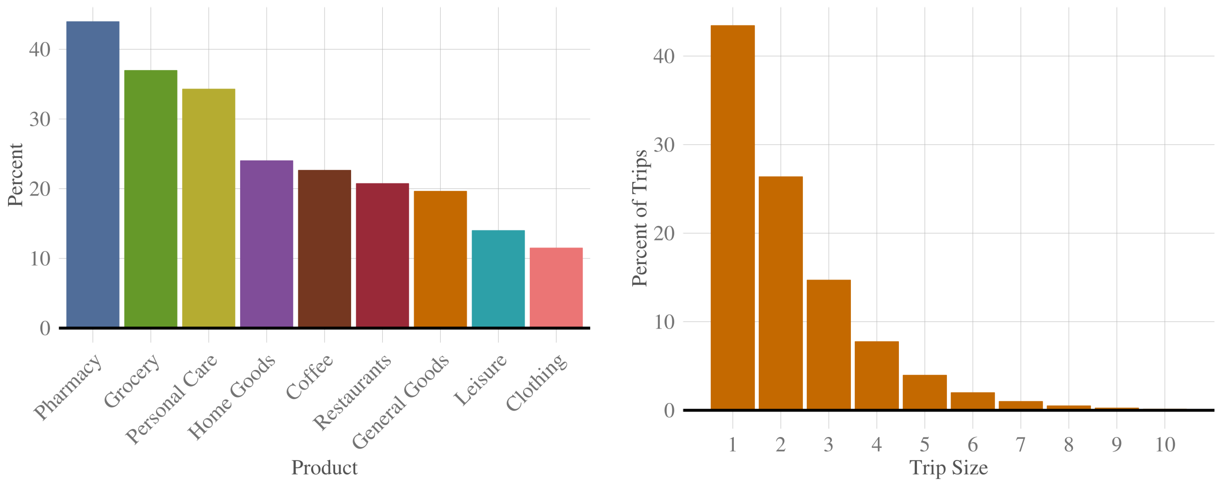
(a) Trips (b) Spending



Notes: Panel (a) shows the percent of days in a month which include an offline purchase for a product. Panel(b) shows the average offline spending per day on each product. General goods include department stores, discount stores, large non-specific retailers, and other miscellaneous retailers like florists and books stores that sell everyday goods. Major categories of personal care services include salons and dry cleaners. Major categories of local leisure include movie theaters and gyms.
Source: Author's calculations using the card transactions from the base 7.7 million customer sample.

Figure A3: Trips Features

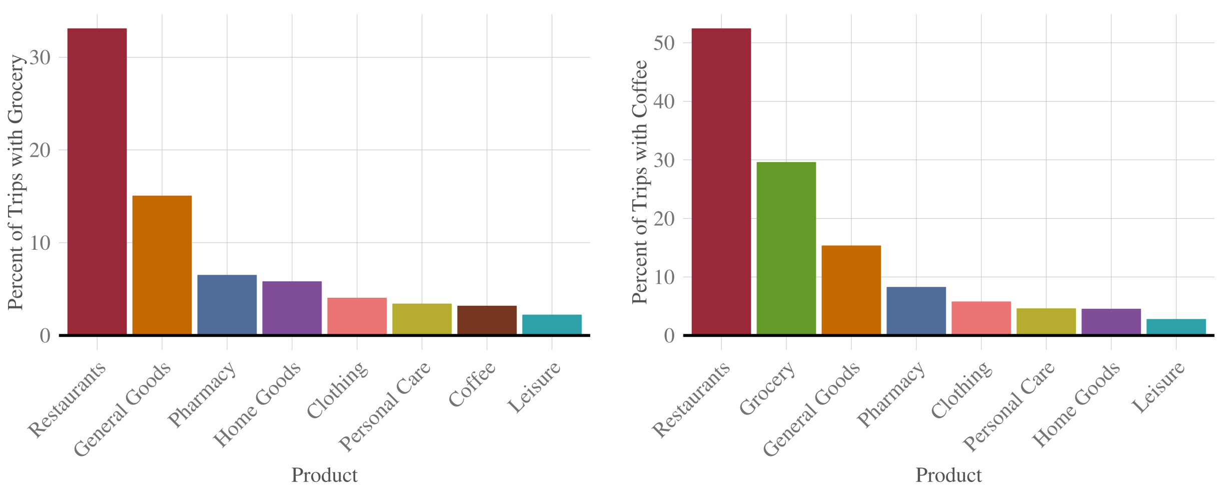
(a) Spending in Home Zip (b) Trip Size Distribution



Notes: Panel (a) shows the share of everyday products purchased in customers' home zip codes. Panel (b) shows the distribution of the number of offline purchases made made on a day with a least one offline purchase.
Source: Author's calculations using the card transactions from the base 7.7 million customer sample.

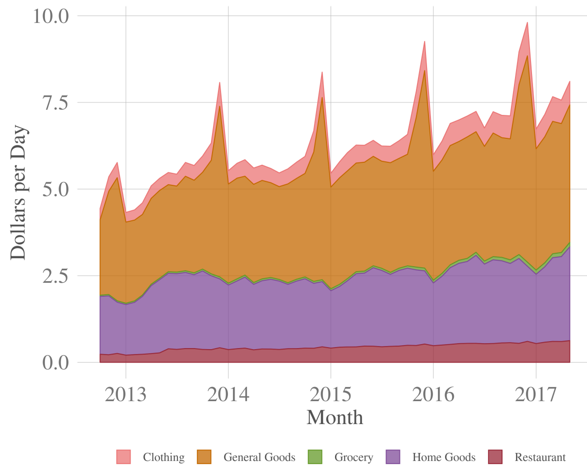
Figure A4: Products Purchased with Grocery and Coffee

(a) Grocery (b) Coffee



Notes: This figure shows the frequency of trips including other everyday products combined with grocery (Panel (a)) and coffee (Panel (b)), assuming customers make at most one offline shopping trip per day.
Source: Author's calculations using the card transactions from the base 7.7 million customer sample.

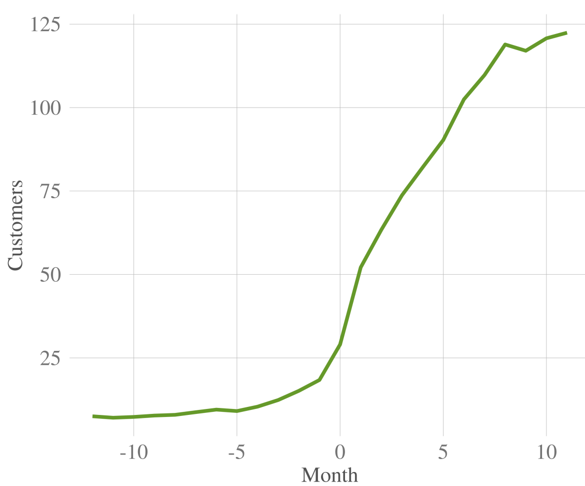
Figure A5: Spending on Online Products



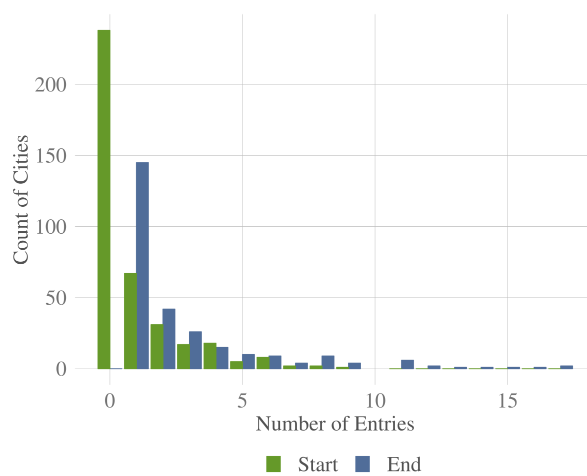
Notes: This figures shows the average online spending per day on each online product. General goods include department stores, discount stores, large non-specific retailers, and other miscellaneous retailers like florists and books stores that sell everyday goods.
Source: Author's calculations using the card transactions from the base 7.7 million customer sample.

Figure A6: Platform Entry

(a) Platform Use at Platform Entry

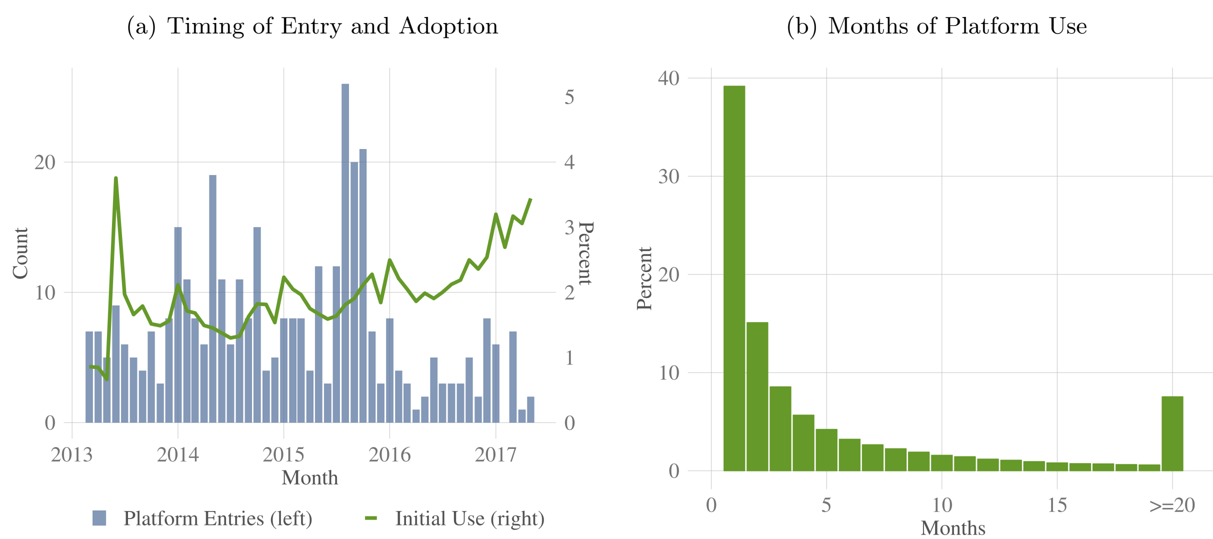


(b) Platforms per City



Notes: Panel (a) shows the average number of customers using a platform in a city before and after the month of entry into that city. The month of entry is estimated to be the first month of "substantial use" from a sequence of months in which customers in the city "continuously" transact with the platform. Substantial use is a threshold defined separately for each platform in each city as a month in which the platform's active customer base in a month is at least 10 percent of the average size of the customer base between the first and last months of observed transaction activity on the platform in the city. This threshold is flexible to allow for varying popularity across platforms and cities. To satisfy continuity, there must exist a sequence larger than 3 months during which transactions are observed in 80 percent of the months. I also require a minimum of two customers transacting with a platform in a city on average between the first and last observed transaction month, inclusive, to qualify as an entry. These choices were calibrated to match entry dates, when known, from publicly available data. Panel (b) shows the number of cities with different numbers of platform entries a the start and end of the sample period among cities with at least one platform by the last month of the panel. There are 17 possible platforms for customers to adopt.
Source: Author's calculations using the 53 billion card transactions of an unbalanced panel of 69 million customers.

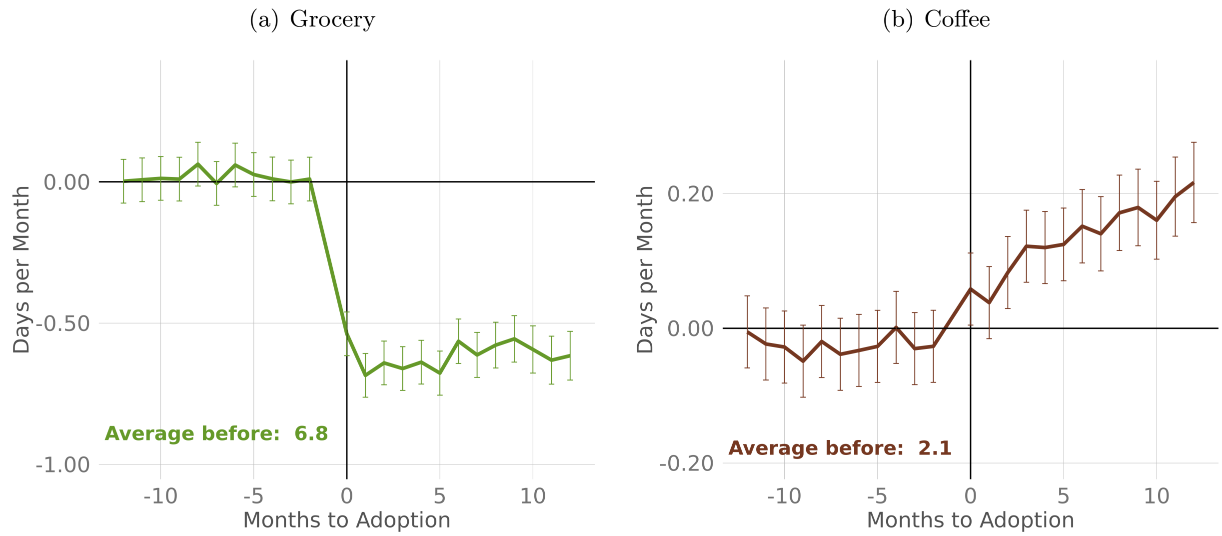
Figure A7: Grocery Platform Adoption Patterns



Notes: Panel (a) shows the time-series of platform entry and the first observed use of an online grocery platform by customers. There is strong seasonality in initial use, with more consumers trying a platform for the first time in the winter months. However, there is no such seasonal pattern to platform entry into cities. Panel (b) shows that more than half of customers that try an online grocery platform only do so for one or two months. However, about one-third use the platform for an extended period of time of five months or more. The latter are considered adopters. There are 17 possible platforms for customers to adopt.
Source: Entry timing is estimated from the 53 billion card transactions of an unbalanced panel of 69 million customers. Adoption statistics are estimated from the 103 thousand customers who use an online grocery platform from the 7.7 million balanced customer panel.

8 Appendix B: Reduced-form Evidence

Figure B1: Trip Effects for Late Adopters



Notes: This figure shows the change in grocery store and coffee trips in the 12 months before and after platform adoption for late adopters of platforms as compared to a matched sample of non-users. In the months after adoption of an online grocery platform, early adopters change their trips comparably to early adopters.
Source: Author's calculations using the transactions of late adopters of online grocery platforms and each of their two nearest neighbors matched on zip code, demographics, and pre-adoption spending patterns.

Table B1: Early Adopter Predictors in October 2014

Category	Covariate	Level		12MΔ		3MΔ	
Gender	Male	-0.334	(0.092)				
	Unknown	-0.446	(0.120)				
Age Bins	Age 1	1.417	(0.870)				
	Age 2	1.350	(0.866)				
	Age 3	0.717	(0.871)				
	Age 4	-0.160	(0.871)				
	Age 5	-0.367	(0.879)				
	Age 6	-1.147	(0.943)				
Income Bins	Income 1	-11.18	(153.989)				
	Income 2	-0.935	(0.184)				
	Income 3	-0.661	(0.165)				
	Income 4	-0.428	(0.152)				
	Income 5	-0.296	(0.155)				
	Income 6	0.081	(0.105)				
Offline Spending	Restaurant	0.0003	(0.004)	-0.048	(0.040)	-0.002	(0.009)
	Grocery	0.020	(0.003)	0.045	(0.039)	0.001	(0.010)
	Leisure	0.0004	(0.020)	-0.022	(0.097)	0.009	(0.025)
	Pharmacy	0.009	(0.016)	-0.031	(0.168)	-0.012	(0.039)
	Personal	0.078	(0.017)	-0.034	(0.128)	0.024	(0.031)
	Coffee	0.152	(0.082)	-0.844	(0.761)	0.108	(0.200)
Offline Trips	Restaurant	0.022	(0.010)	-0.169	(0.114)	0.061	(0.032)
	Leisure	0.022	(0.062)	-0.594	(0.454)	0.102	(0.124)
	Pharmacy	0.010	(0.025)	0.225	(0.266)	0.160	(0.071)
	Personal	-0.016	(0.051)	0.026	(0.463)	-0.168	(0.117)
	Grocery Alone	-0.022	(0.013)	-0.237	(0.162)	0.014	(0.044)
	Coffee Alone	0.019	(0.031)	-0.065	(0.301)	0.024	(0.086)
	Grocery and Coffee	-0.055	(0.055)	0.425	(0.465)	0.049	(0.130)
Online Spending	Restaurant	0.012	(0.007)	0.052	(0.061)	-0.001	(0.014)
	Clothing	-0.011	(0.012)	-0.004	(0.062)	-0.007	(0.013)
	General Goods	-0.003	(0.003)	-0.004	(0.015)	0.008	(0.005)
	Home Goods	-0.0002	(0.001)	-0.006	(0.009)	-0.002	(0.003)
Online Trips	Restaurant	0.093	(0.023)	0.410	(0.278)	0.066	(0.075)
	Clothing	0.054	(0.067)	0.716	(0.487)	-0.121	(0.116)
	General Goods	0.208	(0.016)	0.326	(0.198)	0.045	(0.054)
	Home Goods	0.224	(0.058)	0.227	(0.505)	0.008	(0.136)
Travel Spending	Fuel	-0.037	(0.018)	0.070	(0.183)	0.004	(0.041)
	Transportation	-0.058	(0.016)	0.167	(0.115)	-0.029	(0.016)
Travel Trips	Fuel	-0.018	(0.025)	-0.286	(0.264)	-0.028	(0.069)
	Transportation	0.134	(0.015)	0.406	(0.155)	0.033	(0.041)
Observations	156,766						
Log Likelihood	-3,548.730						

Notes: Standard errors in parentheses. This table shows the correlation between platform adoption and levels and changes in spending and trips for early adopters in October 2014. Changes include 12-month and 3-month periods prior to platform adoption.
Source: Author's calculations using early adopters who adopted in October 2014 and a random sample of non-users.

Table B2: Balance on Covariates Used in Matching

Category	Covariate	*Means*			*t-stats*	
		Adopters	Non-adopters		Non-adopters	
			(All)	(Matched)	(All)	(Matched)
Socioeconomics	Female	0.44	0.37	0.45	15.32	-1.54
	Age Group	2.46	3.16	2.42	-67.43	3.46
	Income Group	2.89	2.81	2.89	3.67	0.2
Offline Spending	Restaurant	15.47	8.16	15.24	58.21	1.4
	Grocery	16.46	10.24	15.52	49.94	4.72
	Leisure	1.08	0.61	1.06	22.89	0.49
	Pharmacy	2.29	1.32	2.2	38.35	2.95
	Personal	2.91	1.12	2.76	53.26	3.34
	Coffee	0.48	0.21	0.48	39.36	0.68
Offline Trips	Restaurant	11.62	8.03	11.8	61.96	-2.45
	Leisure	0.59	0.49	0.59	15.99	0.41
	Pharmacy	2.27	1.55	2.28	40.21	-0.27
	Personal Care	1.31	0.66	1.28	56.27	1.9
	Grocery Alone	6.31	5.86	6.28	12.88	0.65
	Coffee Alone	1.18	0.52	1.2	40.07	-1.2
	Grocery and Coffee	0.49	0.21	0.49	32.2	-0.16
Offline Spending 12MΔ	Restaurant	0.15	0.05	0.16	8.26	-0.23
	Grocery	0.09	0.04	0.07	5.8	1.76
	Leisure	0.02	0	0.02	2.78	-0.19
	Pharmacy	0.01	0	0.01	2.85	0.19
	Personal	0.03	0.01	0.02	4.74	0.94
	Coffee	0.01	0	0	4.43	0.84
Offline Trips 12MΔ	Restaurant	0.05	0.02	0.05	4.9	-0.73
	Leisure	0	0	0	1.92	-0.12
	Pharmacy	0.01	0	0.01	4.25	0.65
	Personal Care	0.01	0	0.01	4.27	0.18
	Grocery Alone	0.01	0.01	0	-2.17	0.26
	Coffee Alone	0.01	0	0.01	2.29	-0.02
	Grocery and Coffee	0	0	0	2.89	0.86
Offline Spending 3MΔ	Restaurant	0.12	0.04	0.14	1.44	-0.25
	Grocery	0.1	0.07	0.09	0.75	0.07
	Leisure	0.04	0	0.03	2.05	0.39
	Pharmacy	0.06	0.01	0.06	3.7	-0.01
	Personal	0.04	0	0.03	2.22	0.59
	Coffee	0	0	0	-0.48	0.03
Offline Trips 3MΔ	Restaurant	0.04	0.01	0.04	2	-0.08
	Leisure	0	0	0	1.76	0.51
	Pharmacy	0.01	0	0.02	2.03	-0.18
	Personal Care	0.01	0	0.01	1.66	-0.7
	Grocery Alone	-0.01	0.02	-0.01	-3.08	-0.11
	Coffee Alone	0	0	0	-0.31	-0.15
	Grocery and Coffee	0	0	0	-0.47	0.85

Notes: This table shows mean differences in spending and trips in levels and changes between early adopters and non-adopters before and after matching. Variables in this table are included in the matching exercise. Levels are 12-month averages prior to adoption and changes are 12- and 3-month changes prior to adoption. For customers who adopt within 12 months of the start of my sample period, the long averages are for the longest time interval available and are at least 6 months. Two-sided t-stats for the differences in means before and after matching are reported.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and a random sample of non-adopters.

Table B3: Balance on Covariates Used in Matching, continued

Category	Covariate	*Means*			*t-stats*	
		Adopters	Non-adopters		Non-adopters	
			(All)	(Matched)	(All)	(Matched)
Online Spending	Restaurant	1.79	0.46	1.63	26.62	2.26
	Clothing	1.96	0.51	1.73	32.36	3.39
	General Goods	9.78	3.11	9.03	46.69	4.24
	Home Goods	5.35	2.03	4.88	17.66	1.81
Online Trips	Restaurant	1.27	0.39	1.23	57.48	1.95
	Clothing	0.58	0.18	0.55	51.32	3.28
	General Goods	3.61	1.27	3.54	79.46	1.66
	Home Goods	0.53	0.22	0.51	44.01	2.35
Online Spending 12MΔ	Restaurant	0.07	0.01	0.05	7.26	1.28
	Clothing	0.04	0.01	0.03	3.22	0.57
	General Goods	0.2	0.03	0.19	9.39	0.46
	Home Goods	0.09	0.01	0.09	2.37	0.16
Online Trips 12MΔ	Restaurant	0.04	0.01	0.04	19.15	0.31
	Clothing	0.02	0	0.01	11.43	1.29
	General Goods	0.09	0.02	0.09	25.95	0.24
	Home Goods	0.01	0	0.01	8.33	0.7
Online Spending 3MΔ	Restaurant	0.14	0.01	0.06	4.21	1.7
	Clothing	0.01	0.02	0.02	-0.18	-0.3
	General Goods	0.31	0.08	0.27	2.91	0.39
	Home Goods	0.15	0	0.28	1.02	-0.72
Online Trips 3MΔ	Restaurant	0.07	0.01	0.06	9.64	0.31
	Clothing	0.02	0.01	0.02	4.37	0.64
	General Goods	0.14	0.04	0.14	10.04	0.27
	Home Goods	0.01	0	0.01	2.48	-0.37
Travel Spending	Fuel	4.62	4.34	4.58	6.93	0.84
	Transportation	2.75	1.09	2.7	47.36	1.15
Travel Trips	Fuel	4.01	4.3	4.01	-8.68	0.05
	Transportation	3.56	1.29	3.52	60.73	0.94
Travel Spending 12MΔ	Fuel	-0.04	-0.04	-0.04	-2.1	-0.38
	Transportation	0.07	0.01	0.06	11.51	0.22
Travel Trips 12MΔ	Fuel	0	0	0	-0.99	-0.58
	Transportation	0.1	0.02	0.09	22.95	0.91
Travel Spending 3MΔ	Fuel	-0.07	-0.05	-0.06	-1.6	-0.39
	Transportation	0.1	0.01	0.09	5.48	0.38
Travel Trips 3MΔ	Fuel	-0.02	-0.01	-0.02	-1.25	0.36
	Transportation	0.14	0.02	0.13	9.99	0.44

Notes: This table shows mean differences in spending and trips in levels and changes between early adopters and non-adopters before and after matching. Variables in this table are included in the matching exercise. Levels are 12-month averages prior to adoption and changes are 12- and 3-month changes prior to adoption. For customers who adopt within 12 months of the start of my sample period, the long averages are for the longest time interval available and are at least 6 months. Two-sided t-stats for the differences in means before and after matching are reported.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and a random sample of non-adopters.

Table B4: Balance on Covariates Not Used in Matching

		Means			*t-stats*	
Category	Covariate	Adopters	Non-adopters		Non-adopters	
			(All)	(Matched)	(All)	(Matched)
Offline Spending	Clothing	4.27	2.21	4.42	29.66	-1.65
	General	7.91	5.31	8.09	23.92	-1.39
	Home	5.29	3.16	5.26	20.71	0.2
Offline Trips	Clothing	1.09	0.83	1.23	28.76	-11.63
	General	3.02	2.66	3.25	19.61	-9.54
	Home	1.13	0.98	1.2	13.87	-5.04
Offline Spending 12MΔ	Clothing	0.02	0	0.01	1.89	0.76
	General	0.03	0	0	1.77	1.8
	Home	0.09	0.01	0.05	3.17	1.29
Offline Trips 12MΔ	Clothing	0	0	0	0.88	-0.34
	General	0	0	0	-0.37	-0.64
	Home	0	0	0	2.08	0.11
Offline Spending 3MΔ	Clothing	0.08	0.03	0.06	1.07	0.38
	General	0.1	0.07	0.03	0.4	0.81
	Home	-0.03	-0.02	-0.08	-0.13	0.31
Offline Trips 3MΔ	Clothing	0.02	0.01	0.01	1.35	0.57
	General	0.02	0.02	0.02	-0.26	-0.34
	Home	0	0	-0.01	0.37	0.74

Notes: This table shows mean differences in spending and trips in levels and changes between early adopters and non-adopters before and after matching. Variables in this table are not included in the matching exercise. Levels are 12-month averages prior to adoption and changes are 12- and 3-month changes prior to adoption. For customers who adopt within 12 months of the start of my sample period, the long averages are for the longest time interval available and are at least 6 months. Two-sided t-stats for the differences in means before and after matching are reported.
Source: Author's calculations using the transactions of 13 thousand early adopters of online grocery platforms and a random sample of non-adopters.

9 Appendix C: Model

9.A Comparison to Logit and Nested Logit Models

The logit model assumes trip utilities are uncorrelated and is equivalent to the PCL model in the event that all $\sigma_{k,l} = 0$. This model provides a simple expression for the log relative probability for the coffee alone trip to the neither trip,

$$ln\Big(\frac{P_{01}}{P_{00}}\Big) = C + \tau ln(2d^c/0.1). \tag{C1}$$

In this expression, the log relative probability only depends on the characteristics specific to those two trips while any characteristic of other trips is irrelevant. In particular, the elasticity of the log relative probability with respect to grocery store value or distances is zero. Therefore, many of the intuitive patterns fundamental to trip choice are not captured. This issue is partially resolved by the nested logit. For example, in the nested logit model in which the two trips containing coffee are nested in one nest and the grocery alone and neither trip are in the other nest, the equivalent expression is

$$ln\Big(\frac{P_{01}}{P_{00}}\Big) = \frac{C + \tau ln(2d^c/0.1)}{1 - \sigma_{01,11}} - \sigma_{01,11} I_{01,11} + \sigma_{00,10} I_{00,10}. \tag{C2}$$

With this structure, features of the grocery alone and chained trips affect elasticities through the inclusive values of the two nests. For example, for a fall in grocery store value,

$$-\frac{\partial ln\Big(\frac{P_{01}}{P_{00}}\Big)}{\partial G} = \frac{\sigma_{01,11}}{1 - \sigma_{01,11}} P_{11|01,11} - \frac{\sigma_{00,10}}{1 - \sigma_{00,10}} P_{10|00,10} \tag{C3}$$

This implies that the elasticity with grocery store value depends on the extent to which the trips within each nest are substitutable with each other and the probability of choosing the trip including the grocery store within each nest. The first remaining issue here is that the nesting choice is arbitrary. These substitution effects only show up because these two trips are in different nests. Second, the remaining irrelevance of alternative assumptions, those that affect intra-nest substitution, continue to create unintuitive patterns. To see this, also consider log probability of coffee alone relative to the chained trip,

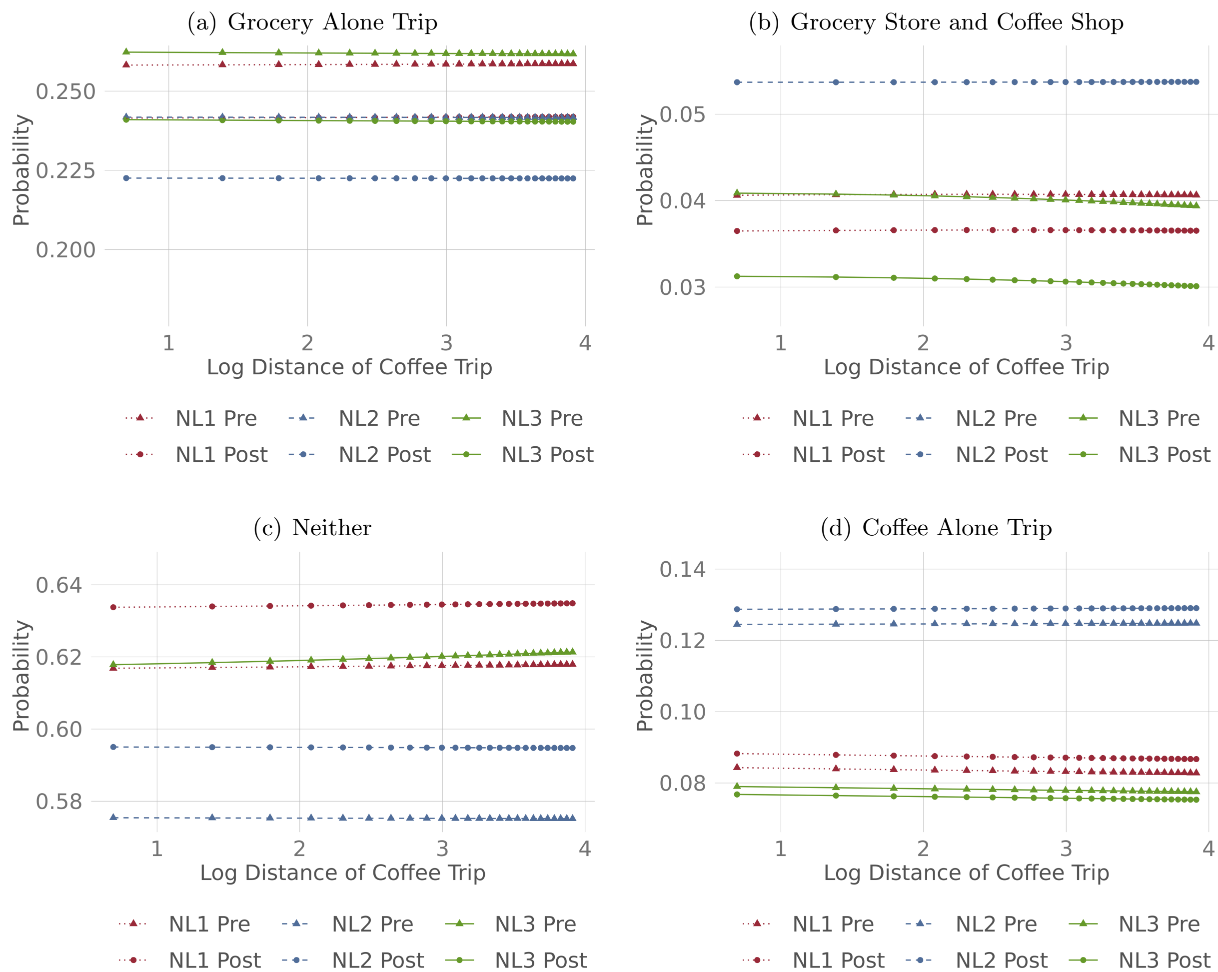
$$ln\Big(\frac{P_{01}}{P_{11}}\Big) = \frac{-G - b + \tau ln(2d^c/(d^g + d^c + d^b))}{1 - \sigma_{01,11}}. \tag{C4}$$

Because the two trips are in the same nest, there are no effects operating through inter-nest substitution. Thus, the elasticity of the relative share of the two trips for a fall in G,

$$-\frac{\partial ln\Big(\frac{P_{01}}{P_{11}}\Big)}{\partial G} = \frac{1}{1 - \sigma_{01,11}}, \tag{C5}$$

is a constant. So while the log relative share of the two trips is affected by relative trip costs, the elasticity of the relative share with respect to a fall in grocery store value is not. This severely limits the ability of the nested logit model to show differentiation in trip substitution affects along attributes, such as distance to the coffee shop in Figure 12, as desired. These restrictions are not limited to one particular choice of nesting structure. As Figure C1 shows, for the simulation exercise in the main text, each of the three possible choices for dividing the four trips into two nests produces patterns in which coffee shop distance has minimal impact on relative substitution patterns.

Figure C1: Substitution Patterns Generated by Nested Logit Models



Notes: This figure shows that each of the three possible versions of the nested logit model produce limited variation in substitution patterns along coffee shop distance. The first nested logit model nests the two trips with coffee into one nest and the two trips without in the other nest. The second nested logit model nests the two trips with grocery into one nest and the two trips without in the other nest. The third model nests the single store trips into one nest and the chained and neither trip into the other nest.

9.B Additional Stores

An extension of the trip choice model beyond the simple one grocery store and one coffee shop setting increases the trip options available to consumers and, therefore, creates more complex substitution patterns. For example, in a setting with just one additional coffee shop, the consumer has six trip options. Denote the additional coffee shop as $c = 2$. The log relative probability of visiting the original coffee shop, $c = 1$ versus no store is as before

$$ln\Big(\frac{P_{01}}{P_{00}}\Big) = ln \sum_{gc' \neq 01} V_{01|01,gc'} - ln \sum_{gc' \neq 00} V_{00|00,gc'} \tag{C6}$$

except that each summation takes place over all five, rather than three, alternative trips pairs. Then, when we calculate the effect of a fall in grocery store values,

$$\begin{aligned} -\frac{\partial ln\Big(\frac{P_{01}}{P_{00}}\Big)}{\partial G} = & \frac{\frac{\sigma_{01,11}}{1-\sigma_{01,11}} P_{11|01,11} V_{01|01,11} + \frac{\sigma_{01,10}}{1-\sigma_{01,10}} P_{10|01,10} V_{01|01,10} + \frac{\sigma_{01,12}}{1-\sigma_{01,12}} P_{12|01,12} V_{01|01,12}}{\sum_{gc' \neq 01} V_{01|01,gc'}} \\ & - \frac{\frac{\sigma_{00,11}}{1-\sigma_{00,11}} P_{11|00,11} V_{00|00,11} + \frac{\sigma_{00,10}}{1-\sigma_{00,10}} P_{10|00,10} V_{00|00,10} + \frac{\sigma_{00,12}}{1-\sigma_{00,12}} P_{12|00,12} V_{00|00,12}}{\sum_{gc' \neq 00} V_{00|00,gc'}}. \end{aligned} \tag{C7}$$

an additional term in the numerator of the first term appears which captures the effect of this fall on the relative attractiveness of a trip to the first coffee shop alone versus the chained trip of the grocery store with the second coffee shop. Similarly, there is an additional term in the numerator of the second term which captures the effect of this fall on the relative attractiveness of the trip to neither store versus the chained trip of the grocery store with the second coffee shop. Thus, this extended model directly captures the effect of losing and winning coffee shops when chains between the grocery store and second coffee shop are broken in favor of trips to the first coffee shop.

Of course, as the model is extended to include more stores, the number of substitutability parameters increase at a rate of 2^N, where N is the size of the trip choice set. To keep the model tractable, it is possible to parameterize substitutability as a function of trip features, such as the type and number of stores included in the trip.

9.C Derivations

In this section, I lay out the derivations used in the discussion of time-use mechanisms in section 4.B. These derivations can also be used to make reduced form predictions and structure alternative estimation procedures, such as non-linear least squares or generalized method of moments. For compactness, define

$$\widetilde{V_{gc}} = \frac{V_{gc}}{(1 - \sigma_{gc,gc'})} \tag{C8}$$

To study the effect of changes in model parameters on trip choices, the effects on $V_{gc|gc,gc'}$, $I_{gc,gc'}$, and $P_{gc}|_{gc,gc'}$ are used. Tables C1, C2, and C3 give the inputs for these for select parameters. The key derivatives for trip utility parameters x and $\sigma_{gc,gc'}$ are:

$$\frac{\partial V_{gc|gc,gc'}}{\partial x} = \Big(\frac{\partial \widetilde{V_{gc}}}{\partial x} - \sigma_{gc,gc'}\frac{\partial I_{gc,gc'}}{\partial x}\Big)V_{gc|gc,gc'} \tag{C9}$$

$$\frac{\partial V_{gc|gc,gc'}}{\partial \sigma_{gc,gc'}} = \Big[\frac{1}{1-\sigma_{gc,gc'}}\widetilde{V_{gc}} - \sigma_{gc,gc'}\frac{\partial I_{gc,gc'}}{\partial \sigma_{gc,gc'}} - I_{gc,gc'}\Big]V_{gc|gc,gc'} \tag{C10}$$

$$\frac{\partial I_{gc,gc'}}{\partial x} = \frac{\partial \widetilde{V_{gc}}}{\partial x} + \Big[\frac{\partial \widetilde{V_{gc'}}}{\partial x} - \frac{\partial \widetilde{V_{gc}}}{\partial x}\Big]P_{gc'}|_{gc,gc'} \tag{C11}$$

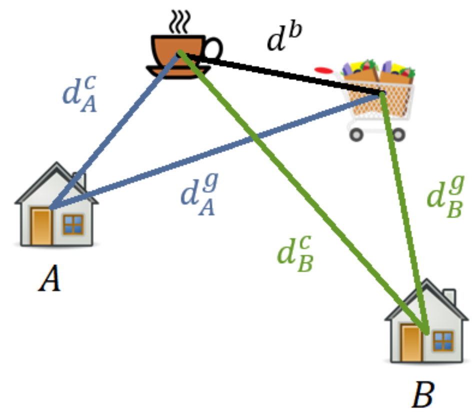
$$\frac{\partial I_{gc,gc'}}{\partial \sigma_{gc,gc'}} = \frac{1}{1-\sigma_{gc,gc'}}\Big[\widetilde{V_{gc}}P_{gc}|_{gc,gc'} + \widetilde{V_{gc'}}P_{gc'}|_{gc,gc'}\Big] \tag{C12}$$

$$\frac{\partial P_{gc}|_{gc,gc'}}{\partial x} = \Big[\frac{\partial \widetilde{V_{gc}}}{\partial x} - \frac{\partial \widetilde{V_{gc'}}}{\partial x}\Big]P_{gc}|_{gc,gc'}P_{gc'}|_{gc,gc'} \tag{C13}$$

$$\frac{\partial P_{gc}|_{gc,gc'}}{\partial \sigma_{gc,gc'}} = \frac{1}{1-\sigma_{gc,gc'}}P_{gc}|_{gc,gc'}P_{gc'}|_{gc,gc'}(\widetilde{V_{gc}} - \widetilde{V_{gc'}}) \tag{C14}$$

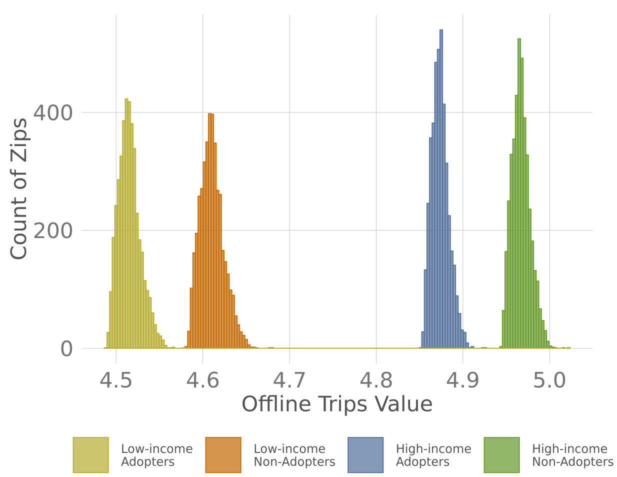
9.D Figures and Tables

Figure C2: Consumer Travel Costs



Notes: This figure illustrates that each consumer faces unique travel costs for each trip because she lives in her own location with its unique relative distances to all the stores.

Figure C3: Trip Value Distributions



Notes: This figure shows the distribution of offline trip values for low- and high-income consumers pre- and post-online grocery platform adoption across zip codes. Values for low-income consumers are lower than high-income consumers because they value grocery stores and coffee shops less. The distributions are also narrower because they have a smaller disutility for distance.
Source: Author's calculations using zip codes with at least 500 customers from the panel and positive adoption rates for both low- and high-income consumers. There are 2,276 such zip codes.

Table C1: Trip Value Derivatives

	d^g	d^c	d^b	b	τ
V_{10}	$\frac{\tau}{d^g}$	0	0	0	$ln(2d^g)$
V_{11}	$\frac{\tau}{d^g+d^c+d^b}$	$\frac{\tau}{d^g+d^c+d^b}$	$\frac{\tau}{d^g+d^c+d^b}$	1	$ln(d^g + d^c + d^b)$
V_{01}	0	$\frac{\tau}{d^c}$	0	0	$ln(2d^c)$
V_{00}	0	0	0	0	$ln(0.1)$

Notes: This table shows the derivatives for each trip value with respect to the parameter in the column heading.

Table C2: Inclusive Value Derivates

	d^g	d^c	d^b	b	τ
$I_{01,00}$	0	$\frac{\tau}{d^c} P_{01\|00,01}$	0	0	$ln\left(2d^c\right) + ln\left(\frac{0.1}{2d^c}\right) P_{00\|01,11}$
$I_{01,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|01,11}$	$\frac{\tau}{d^c} + \left[\frac{\tau}{d^g+d^c+d^b} - \frac{\tau}{d^c}\right] P_{11\|01,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|01,11}$	$P_{11\|01,11}$	$ln\left(2d^c\right) + ln\left(\frac{d^g+d^c+d^b}{2d^c}\right) P_{11\|01,11}$
$I_{01,10}$	$\frac{\tau}{d^g} P_{10\|10,01}$	$\frac{\tau}{d^c} P_{01\|10,01}$	0	0	$ln\left(2d^c\right) + ln\left(\frac{d^g}{d^c}\right) P_{10\|01,10}$
$I_{10,00}$	$\frac{\tau}{d^g} P_{10\|01,00}$	0	0	0	$ln\left(2d^g\right) + ln\left(\frac{0.1}{2d^g}\right) P_{00\|01,11}$
$I_{10,11}$	$\frac{\tau}{d^g} + \left[\frac{\tau}{d^g+d^c+d^b} - \frac{\tau}{d^g}\right] P_{11\|10,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|10,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|10,11}$	$P_{11\|10,11}$	$ln\left(2d^g\right) + ln\left(\frac{d^g+d^c+d^b}{2d^g}\right) P_{11\|10,11}$
$I_{00,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|00,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|00,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|00,11}$	$P_{11\|00,11}$	$ln\left(d^g + d^c + d^b\right) + ln\left(\frac{0.1}{d^g+d^c+d^b}\right) P_{11\|00,11}$

Notes: This tables shows the derivates for each inclusive value with respect to the parameter in the column-heading. Each cell should be multiplied by the appropriate $1/(1-\sigma_{gc,gc'})$.

Table C3: Conditional Probability Derivatives

	d^g	d^c	d^b	τ	b
$P_{11\|01,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|01,11} P_{01\|01,11}$	$\left[\frac{\tau}{d^g+d^c+d^b} - \frac{\tau}{d^c}\right] P_{11\|01,11} P_{01\|01,11}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|01,11} P_{01\|01,11}$	$ln(\frac{d^g+d^c+d^b}{2d^c}) P_{11\|01,11} P_{01\|01,11}$	$P_{11\|01,11} P_{01\|01,11}$
$P_{10\|00,10}$	$\frac{\tau}{d^g} P_{10\|00,10} P_{00\|00,10}$	0	0	$ln(\frac{2d^g}{0.1}) P_{10\|00,10} P_{00\|00,10}$	0
$P_{11\|11,10}$	$\left[\frac{\tau}{d^g+d^c+d^b} - \frac{\tau}{d^g}\right] P_{10\|11,10} P_{11\|11,10}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|11,10} P_{10\|11,10}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|11,10} P_{10\|11,10}$	$ln(\frac{d^g+d^c+d^b}{2d^g}) P_{11\|11,10} P_{10\|11,10}$	$P_{11\|11,10} P_{10\|11,10}$
$P_{10\|11,10}$	$\left[\frac{\tau}{d^g} - \frac{\tau}{d^g+d^c+d^b}\right] P_{10\|11,10} P_{11\|11,10}$	$-\frac{\tau}{d^g+d^c+d^b} P_{10\|11,10} P_{11\|11,10}$	$-\frac{\tau}{d^g+d^c+d^b} P_{10\|11,10} P_{11\|11,10}$	$ln(\frac{2d^g}{d^g+d^c+d^b}) P_{10\|11,10} P_{11\|11,10}$	$-P_{10\|11,10} P_{11\|11,10}$
$P_{10\|01,10}$	$\left[\frac{\tau}{d^g} - \frac{\tau}{d^c}\right] P_{10\|01,10} P_{01\|01,10}$	$-\frac{\tau}{d^c} P_{10\|01,10} P_{01\|01,10}$	0	$ln(\frac{d^g}{d^c}) P_{10\|01,10} P_{01\|01,10}$	0
$P_{11\|11,00}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|11,00} P_{00\|11,00}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|11,00} P_{00\|11,00}$	$\frac{\tau}{d^g+d^c+d^b} P_{11\|11,00} P_{00\|11,00}$	$ln(\frac{d^g+d^c+d^b}{0.1}) P_{11\|11,00} P_{00\|11,00}$	$P_{11\|11,00} P_{00\|11,00}$

Notes: This tables shows the derivates for each conditional probability with respect to the parameter in the column heading. Each cell should be multiplied by the appropriate $1/(1-\sigma_{gc,gc'})$.

Table C4: MLE Estimation Results

	Model		
	PCL	NL1	Logit
[1] Both:(intercept)	-1.487	-2.258	-2.812
	(0.053)	(0.079)	(0.011)
[2] Coffee:(intercept)	-1.013	-1.772	-1.828
	(0.009)	(0.027)	(0.00)
[3] Grocery:(intercept)	-0.994	-0.827	-1.013
	(0.010)	(0.081)	(0.007)
[4] τ_l	-0.017	-0.034	-0.049
	(0.001)	(0.003)	(0.002)
[5] τ_h	-0.010	-0.005	-0.012
	(0.001)	(0.001)	(0.001)
[6] Both:HI	0.242	0.060	0.152
	(0.023)	(0.014)	(0.012)
[7] Coffee:HI	0.197	-0.072	-0.128
	(0.008)	(0.012)	(0.010)
[8] Grocery:HI	0.235	0.148	0.192
	(0.007)	(0.017)	(0.008)
[9] Both:Post:EA	-0.111	-0.063	-0.129
	(0.012)	(0.018)	(0.023)
[10] Coffee:Post:EA	-0.091	0.007	-0.020
	(0.008)	(0.013)	(0.014)
[11] Grocery:Post:EA	-0.107	-0.115	-0.150
	(0.008)	(0.013)	(0.00)
[12] Both:Post:EA:HI	0.052	0.041	0.066
	(0.012)	(0.017)	(0.024)
[13] Coffee:Post:EA:HI	0.050	0.023	0.022
	(0.009)	(0.013)	(0.015)
[14] Grocery:Post:EA:HI	0.038	0.039	0.056
	(0.009)	(0.009)	(0.010)
[15] $1-\sigma_{00,11}$	0.243		
	(0.185)		
[16] $1-\sigma_{01,00}$	0.091		
	(0.746)		
[17] $1-\sigma_{01,11}$	0.435	0.495	
	(0.083)	(0.070)	
[18] $1-\sigma_{10,11}$	0.267		
	(0.053)		
[19] $1-\sigma_{01,10}$	0.026		
	(0.001)		
[20] $1-\sigma_{00,10}$		0.806	
		(0.076)	
Post	Y	Y	Y
EA	Y	Y	Y
Post:HI	Y	Y	Y
EA:HI	Y	Y	Y
Observations	8,028,341	8,028,341	8,028,341
R^2	0.002	0.002	0.002
Log Likelihood	-7,629,270.000	-7,630,291.000	-7,630,012.000

Notes: Standard errors in parentheses. This table shows model estimates from the maximum likelihood estimation of equation 16 that map to Table 1. Grocery store and coffee shop values for low-income consumers are the intercepts, coefficeints [2] and [3]. The fixed benefit for low-income consumers is the difference in the Both intercept, coefficient [1], and the combined store value implied by the sum of [2] and [3]. Similar logic holds for high-income consumers and post-adoption effects on grocery store value.
Source: Authors calculations from 8,604 early adopters and matched controls who meet these requirements.