Buck Wild: The Impact of the Dollar Store on Households and Local Retail Competition[*]

Camilla Schneier and Yixin Sun

May 25, 2023

Abstract

Dollar stores have dramatically changed the food retail landscape of the United States. This paper estimates the effects of the dollar store on household welfare and local retail competition. Leveraging the first dollar store entry into a zip code, we show that dollar store entry expands the set of goods purchased by households without significantly affecting local retail competition, thus increasing household welfare. As a result of this expanded choice set, we show that households reduce total food expenditures, quantities, and the number of unique products consumed. We show that dollar stores' introduction of lower-priced goods (compared to preexisting grocery alternatives) shifts household consumption towards bundles with lower prices but fewer varieties. In a model of household consumption, we estimate the value of novel dollar store entry to the household at 12% of annual grocery expenditure.

[*]Schneier: cschneier@uchicago.edu, University of Chicago; Sun: ysun9@chicagobooth.edu, University of Chicago Booth School of Business. We thank Fern Ramoutar for conceptualizing the use of retail closings as an instrument. We are grateful to Milena Almagro, Joe Battles, Marianne Bertrand, Eric Budish, Neil Cholli, Thomas Covert, Michael Dinerstein, Jean-Pierre Dubé, Greg Kaplan, Michael Galperin, Mikhail Golosov, Veronica Guerrieri, Matthew Notowidigdo, Fern Ramoutar, Francesco Ruggieri, Alexander Torgovitsky and participants at the Food Access at Dollar Stores (FADS) workshop, University of Chicago Economics Third Year Seminar, University of Chicago IO working group, and University of Chicago Public/Labor seminar for helpful comments on the paper. Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. This research was supported in part through the computational resources and staff contributions provided for the Mercury high performance computing cluster at The University of Chicago Booth School of Business which is supported by the Office of the Dean. This research is funded by the Becker Friedman Institute at the University of Chicago. Schneier gratefully acknowledges support from the National Science Foundation Graduate Research Fellowship.

1 Introduction

The dollar store has emerged as an important food source for many Americans. With over 35,000 stores in America – more than there are Walmart, Starbucks, and McDonald's combined, and having opened half of new retail stores in the last year, the dollar store format has garnered simultaneous praise and concern. In the best case, the dollar store unilaterally increases supply, and the increase in supply translates to more choices for consumers and higher consumer welfare. In the worst case, the supply-side response from other retailers leads to a net reduction in supply which shrinks the household choice set and decreases household welfare. In this scenario, grocery stores and other retailers are out competed and exit the market, creating food deserts.[1] Overall, the debate centers on the effect of the dollar store on consumer behavior and local competition, and the implication for consumer welfare.

This paper quantifies the impact of the dollar store format on consumer welfare and assesses which aspects of the dollar store bundle have the biggest effects. We focus on the first dollar store entry in a zip code and analyze the effect on prices, quantities, and varieties of goods purchased. First, we document the retail environment post-dollar store entry and the consumer response to the changed retail environment. We find that the dollar store unilaterally increases the household choice set, and the average household reacts to the dollar store by reducing the number of unique bar codes it consumes by 6% and the quantity of dry goods it consumes by 4%. Specifically, the household consumption bundle shifts from a larger bundle with more variety and higher prices towards a smaller bundle with lower prices but fewer varieties. As a result, the average household total food expenditure declines by 5%. Next, we investigate which features of the dollar store drive this effect. Our demand estimates indicate that households re-optimize in this way solely because the dollar store provides low-cost goods. We quantify value of the dollar store and find that the average household values the dollar store at 12% of annual grocery expenditure.

To understand how the dollar store changes the household choice set, we establish stylized facts on the the types of goods the dollar store provides. We find that dollar stores offer much lower prices, fewer varieties, and largely dry goods. First, for the same product, dollar stores are much cheaper than all other store types: compared to grocery stores, we find that dollar stores consumers pay 23% less for the same product, even after controlling for the size of the good. Second, dollar stores have much less variety in their offerings at every level; compared to grocery stores, dollar stores have 20% fewer unique bar codes. Third, we find that dollar stores specialize heavily in dry goods, and offer a different product assortment than grocery stores.

To understand how consumers and retailers react to the dollar store, we employ an event study model of the first dollar store entry into a zip code. For households, we study how the dollar store affects expenditures, prices, quantities, varieties, and shopping trips to other retailers. For retailers, we study the effect on prices and grocery store exit. We find that the first dollar store entry results in a 5% decrease in household food expenditures. These household savings are driven by substitution from the grocery store and the discount store to the dollar store, a

[1] Critics also worry that the dollar store format is nominally cheap but expensive per unit and unhealthy and shifts consumers away from fresh produce.

reduction in quantity, and an overall decrease in consumption variety. We find that the dollar store effect is persistent, lasting for at least five years after the first entry.

For the event study, we focus on the first dollar store entry in a household's zip code and our control group are not-yet-treated households. We choose these specifications because, first, the first dollar store most cleanly captures the effect of the dollar store format on consumers: focusing on further entry would instead capture switches across dollar stores as well as the effect on new dollar store customers.[2] Second, 60% of zip codes with at least one dollar store have only one dollar store (40% of population weighted zip codes). Thus, the first dollar store entry captures the modal entry effect.[3] Third, 70% of zip codes have no dollar store (55% of population weighted zip codes), and so the first dollar store is the most relevant margin of entry for the vast majority of zip codes. Finally, our control group constitutes households in not-yet-treated zip codes in order to account for the fact that some of the untreated zip codes will never get a dollar store precisely because they are not comparable to zip codes that do get dollar stores.

To ensure we have isolated the effect of the dollar format on consumers, we rule out other potentially important channels through which the dollar store could have reduced household expenditure. We focus on economically plausible and relevant channels that have been discussed in related retailer entry literature, as well as those highlighted in policy and media discussions: incumbent retailer response and other changes in household behavior. On the incumbent side, first we show that the dollar store does not induce a price response from other retailers. Second, we show that on average, grocery stores do not exit following the first dollar store entry. Third, we show that the average household continues taking the same number of trips to other store types after the first dollar store enters. Given these three facts, we reason that first dollar store does not preclude shopping at other store types and the first dollar store entry is not accompanied by a decline in the household choice set.

Mechanically, three channels can contribute to the drop in expenditure: (1) product choices stay the same, price decreases (2) product choices stay the same, but quantity decreases, and (3) household product choices change. We rule out the price channel; for the same consumption bundle, households do not pay lower prices than the state average. Rather, we observe that households are re-optimizing product choices such that their consumption bundle shifts from a larger bundle with more variety and higher prices towards a smaller bundle with lower prices but fewer varieties. Households reduce quantity, as measured in ounces[4], indicating that prior to dollar store entry, households lacked smaller-sized options and thus purchased more than the optimal amount of goods. The results suggest that the dollar store offers new price-size-variety combinations.

We model consumer choice in order to understand how dollar store products substitute for

[2] We study the change in shopping patterns as a result of the dollar store format, not the change between dollar stores. However, we note that some households will have previously shopped at dollar stores in zip codes other than their own.

[3] As shown in Figure A.3, zip codes with two dollar store constitute less than 8% of zip codes, zip codes with three dollar stores constitute less than 5% of zip codes, and there are even fewer zip codes with more dollar stores.

[4] Ounces are a unit of weight popular in the United States.

products at other store types and to understand how consumers value different aspects of the dollar store bundle. Specifically, we model consumer choice as a nested logit model wherein a household chooses a product group from a store type. We choose to model this level of aggregation because policy concerns focus on whether the dollar store induces consumers to substitute across these large product categories (for example, from fresh produce at the grocery store to snacks at the dollar store). We find that broad product groups are much more substitutable within store types than across store types.

Furthermore, the demand estimates show which of the dollar store characteristics (low prices, fewer varieties, smaller sizes, and a different composition of goods) drive the change in the household consumption bundle. We find that households prefer lower prices, larger package sizes, and more variety, indicating that on these dimensions, the dollar store's main benefit is its low price point. Since dollar stores carry a limited selection of goods, we use the product group fixed effect estimates to determine how the dollar store selection compares to a selection of goods with similar number of product groups but based entirely on consumer preferences. We find that the dollar store offers different goods than those most favored by consumers. That is, the dollar stores' unique advantage over other store types is not the size, variety, or product groups sold; the dollar stores' unique advantage rests entirely on its ability to provide its' consumers with low prices. Under this model, low prices are the demand-side driver of dollar stores' successful proliferation.

Since prices and quantities are determined in equilibrium, we introduce a novel instrument for the demand estimation which exploits a plausibly exogenous change in the cost for the dollar store (and other retailers) to enter a neighborhood. Throughout the 2010s, several non-food brick and mortar stores (e.g., Blockbuster) went bankrupt, and these national-level bankruptcies forced these retailers to close their stores in a short time frame. These national bankruptcy-induced closures were independent of the local demand for food. However, these store closings lowered the cost for new food retailers to enter the zip code. We leverage these supply-side shocks as exogenous variation in the supply of food.

To quantify the value of the dollar store, we compute the compensating variation, the compensation required for a household without a dollar store to be indifferent to a household with a dollar store. We find that the average household would have to be compensated 12% of yearly grocery expenditure.

This work addresses address several of the policy concerns surrounding the effect of dollar stores on consumers and local retail competition. First, there is a concern that dollar store goods are cheaper only because they are smaller, but are not cheaper on a per-unit basis. Our descriptive results show that across almost all food groups, dollar store goods are the cheapest on a per unit basis. Second, we address the policy concern that the dollar store causes people to shift away from healthy foods at the grocery store toward unhealthy foods at the dollar stores . With our event study results, we show that the quantity (in ounces) of fresh produce does not change following the first dollar store entry. With our demand estimates, we show that broad product groups within store types are much closer substitutes than broad product groups across store types. Third, our results corroborate the understanding that the grocery store is essential to households, as shown in Allcott et al. (2019), as our event study results show that even as a new food-providing retailer enters the household's zip code, the number of trips to the grocery store stays the same. Fourth, we study the effect of the first dollar store on the household choice set, showing it expands. Fifth, we quantify the welfare impacts of the dollar store.

Finally, we turn our attention to the heterogeneous impact of the dollar store. We show that dollar stores locate disproportionately in lower-income neighborhoods and sparse retail environments (we consider non-metro neighborhoods and food deserts). We re-run our main analysis for different income groups, as well as for households in sparse retail environments. We find that although lower-income groups and households in sparse retail environments tend to shop more on average at the dollar store, the patterns – reductions in average household total food expenditure and a decline in unique bar code variety – largely hold across all groups.

1.1 Related literature

Our work contributes to three strands of literature. First, there is extensive work on the entry of big-box retailers (e.g. Walmart, K-mart, Sam's Club) on consumer welfare and local competition. The existing literature finds that big-box stores increase consumer welfare by offering substantially lower prices, while simultaneously driving down revenue at local incumbent stores.[5] Distinguishing our research from the previous literature on entry of big-box stores, we find that dollar store format is distinct in its product offerings, prices, and sizes. Unlike with the big-box format, we find that the first dollar store adds low-priced varieties to the market without inducing grocery exit or price competition, and thus is unambiguously welfare increasing.

Second, this paper adds to the nascent literature on the role of dollar stores in the US economy. As with the big box store literature, this literature focuses on how dollar stores compete with preexisting retailers (a large focus is on competition with grocery stores), and how dollar stores influence consumer shopping behavior. In particular, studies have found that dollar stores compete with and cause some grocery store exit, particularly for independent grocers (see, for example, Chenarides et al. (2021) and Caoui et al. (2022)).[6] With the goal of understanding the effect of the dollar store format on the local retail environment, we study the effect of the first dollar store entry on prices and exits. On average, we find a precise null price response and a lack of grocery store exit.[7] Given the lack of supply side response, we interpret the

[5]For work on the impact of discount stores on consumer welfare, see Hausman and Leibtag (2007), Basker and Noel (2009), Ailawadi et al. (2010), Arcidiacono et al. (2019), and Atkin et al. (2018). For work on retail competition, see Hausman and Leibtag (2007), Basker and Noel (2009), Ailawadi et al. (2010), Arcidiacono et al. (2019), Atkin et al. (2018), Leung and Li (2021)), Jia (2008), Ellickson and Grieco (2013), and Arcidiacono et al. (2016). Meanwhile, the effect of big-box retailers on incumbent stores' prices is ambiguous; for example, a recent paper by Arcidiacono et al. (2019) that challenges previous findings of price competition using new event study methodology.

[6]In more depth: on spatial competition, Chenarides et al. (2021)'s model conclude that dollar stores directly compete with other stores of similar size and format, but complement stores of different formats. Caoui et al. (2022) model a dynamic game of dollar store entry, exit, and investment played between dollar store chains and independent grocers, and find dollar stores out-compete independent grocery stores.

[7]This result is not inconsistent with the literature; for example, Caoui et al. (2022) find a 7% decrease in independent grocery stores following the first dollar store entry, which is consistent with zero exit on average

consumer response as directly driven by the dollar store format (as opposed to also driven by local retailer's response to the dollar store entry).

The literature has also begun to study the effect of the dollar store on household shopping patterns and welfare. The literature has documented that dollar store entry causes a decline in household total food expenditure. Studying the consumer response, we find that the first dollar store expands the household choice set and pushes households towards a bundle with lower prices but fewer quantities and varieties. On consumer welfare, Cao (2022) find that the dollar store increases welfare, focusing on the provision of private-label products. Since we show that the household choice set responds, we find that household welfare increases, and quantify it as 12% of annual grocery store expenditures. We also use our model to address the open question as to how the dollar stores proliferated so effectively: according to the demand estimates, the dollar store's core value to its consumers are its provision of low-cost goods.[8][9]

Third, we contribute to the broader literature on the effect of supply-side retailers on consumer shopping behavior. One strand of this literature seeks to understand the strong correlation between neighborhood income and availability of healthy foods. Allcott et al. (2019) use grocery entry to study the impacts of access on nutritional inequality and identify that 90% of the difference in nutritional inequality is driven by demand side differences, while only 10% are driven by food access and prices. Ex ante, we might expect that dollar stores play a different role in nutritional inequality. For example, dollar stores are hypothesized to thrive and compete in low income areas and sparse retail environments. However, we also find dollar store entry does not change the number of trips to the grocery store, and corroborate the result that households will travel to the grocery store regardless of other alternatives. In a second vein, our paper adds to the literature that documents how expansions of the household choice set can result in decreases in consumed product variety and a null price response (see Illanes and Moshary (2020) for this in the context of the liquor market or Natan (2020) in the context of takeout restaurants). Methods-wise, we add a novel cost-shifting instrument to estimate demand.

considering all grocery stores.

[8] We discuss supply-side drivers such as store size and leasing strategy in the context of our instrument, and for more on the supply side, such as use of the dollar store distribution system, see Caoui et al. (2022)).

[9] There are a few papers that help characterize the important features of the dollar store. For example, Chenarides et al. (2021) provide descriptive statistics on locations of dollar stores, finding that dollar stores are most prevalent in non-metro areas, and also establish that correlationally, once a dollar store enters a food desert, that area is more likely to remain a food desert. Shannon (2021) show that dollar stores exist disproportionately in minority neighborhoods. Schmall et al. (2021) survey dollar store shoppers to find that dollar store shoppers enjoyed dollar stores' low prices, convenience, and variety. Feng et al. (2023) study how dollar store expenditure shares and dollar store expenditure growth compares with other food types. We also detail the ways in which dollar store prices are cheaper, sizes are smaller, and varieties within product groups are far less compared to other retailers. In particular, dollar store price, variety, and size discounts relative to the grocery store vary significantly by good, which is unique to the dollar store store type. In this way, although they tend to offer the same product groups, dollar stores are very different from discount stores.

2 Data

We use Homescan (HMS) Nielsen data between 2008 and 2019 to study the effect of the dollar store on consumers. The HMS data tracks 40,000 to 60,000 US households and their retail purchases. HMS households scan UPCs of all consumer packaged goods they purchase from any store. In addition to transaction variables, HMS also reports demographic variables such as household income, household composition, household size, number of children, race, and the age, education, employment status, and hours work for male and female household heads. Nielsen tracks the store type (called channel type) households shop at, allowing us to identify purchases specific to dollar stores, grocery stores, discount stores, etc. For our analysis, we follow Nielsen's definition of store types (dollar stores, grocery stores, club stores, convenience stores, drug stores, and discount stores).

Because many of most salient policy questions are around food tradeoffs, we study the effect of dollar stores on food purchasing behavior. Figure 1 details the food terminology in the Nielsen HMS household transactions dataset (figure source: Handbury (2021)). A UPC/barcode uniquely identifies a product, such as "Campbell's Beef Stew", which has a different code than "Campbell's Chicken Noodle Soup". Products with the same UPC are identical in composition, size, and brand. One level up is what Nielsen calls a "product module", which describes the type of good the UPC/barcode belongs to. Products in the same product module may have different sizes and brands but are very similar within the packaging. One more level up are product groups: product groups contain products that are not identical, but broadly similar (for example "Fresh Produce" and "Snacks" are both product groups). In this example, Campbell's Beef Stew would belong to the product module "Stew - Beef - Shelf Stable", which belongs to the product group "Prepared Food - Ready-to-Serve". Policy concerns of how dollar stores induces changes in shopping behavior have focused on dollar store-induced substitution across product groups. At the highest level of aggregation, Nielsen defines six food departments (dairy, deli, dry grocery, fresh produce, frozen foods, and packaged meats).[10]

Figure 1: Nielsen Definitions

UPC/Barcode		Product Module		Product Group
(ex: Campbell's Beef Stew)	$\subset$	(ex: Stew - Beef - Shelf Stable)	$\subset$	(ex: Prepared Food - Ready-to-Serve)
n = 2,152,647		n = 856		n = 62

For each UPC/barcode bought, we compute price (per good) as the total price paid minus the coupon value, divided by the quantity of the UPC/barcode purchased. In order to compare across similar products with different package sizes, we compute the price per unit as the price per good divided by the amount (or size) of the product. We focus on food purchases.[11] We define quantities as the amount or the weight in ounces of each good.[12] For the event study,

[10]For fresh produce, department and product group are the same.

[11]Furthermore, we eliminate magnet data from the sample.

[12]Nielsen provides weight data in ounces (weight), pounds (weight), fluid ounces (volume), quarts (volume), and counts (dimensionless). Because they are the most popular units, we use ounces for our quantity for dry

the panel is balanced by restricting the sample to any household satisfying two criteria. First, the household is in a zip code with an eventual dollar store entry. Second, the household is observed in the same zip code in the data in the year before and after the dollar store enters.

For demand estimation, we define a market as a set of product groups in stores in a county in a year for each income rank (we follow Allcott et al. (2019)).[13] We subset to the set of households in our event study to keep the underlying data comparable and consistent throughout the paper. That is, we require a balanced panel and only include households in zip codes that will eventually receive a dollar store. We split our households into four groups by income rank and compute separate demand parameters for each income group, as income is an important dimension of heterogeneity (dollar stores disproportionately exist in low income neighborhoods), different income groups are expected to have different price elasticities, and because this is standard practice (for example, see Allcott et al. (2019), Atkin et al. (2018)). In addition, we drop observations whenever the price for that good is zero.

We use dollar store locations compiled from a database of Supplemental Nutrition Assistance Program (SNAP) authorized retailers. This data spans 1990-2019, and records the date and location of a store when it enters the SNAP database. Since the majority of locations from the major dollar store chains had become SNAP retailers by 2008, and new stores after that period are likely to automatically enroll in SNAP, we focus on the time period between 2008 and 2019. We subset our SNAP data to the five biggest dollar store chains: Family Dollar, Dollar Tree, Dollar General, Fred's, and 99 Cents Only (roughly 85% of all dollar stores).[14]

For grocery store counts in each zip by year we use the ZIP Code Business Patterns (ZBPs) as in Allcott et al. (2019).

The retail closing instrument is constructed with data from Infogroup, which provides a historical, yearly directory information for U.S. companies, including name, address, estimated sales and number of employees. Specifically, we compute the number of retailers in each zip in each year for the most popular non-food retailers that went bankrupt throughout the 2010s. In Table A.2 we include a list of retailers and their bankruptcy year, and in Figure A.19 we plot the number of stores in each year.

goods and fluid ounces for liquid goods, converting measurements in pounds and quarts. We impute weight measured in counts whenever possible, and eliminate data when we cannot impute a weight. The list of imputed weights is found in Table A.5.

[13] To estimate demand, we use a county to define a market – as opposed to a zip code, which we used in the event study analysis. In the event study analysis, we consider the effect on the households located nearest to the dollar store. However, the market for food is likely larger than a zip code, especially for grocery stores.

[14] As the SNAP dataset is not standardized, finding all true dollar stores outside of these chains would be extremely burdensome. In addition, these five chains make up roughly 85% of all dollar stores, and we are thus able to capture most of the market with these.

3 Descriptive Statistics - What is a Dollar Store?

In this section, we describe what makes a dollar store unique to other store types. Dollar stores are conceptualized by a single price point (1$). And while dollar stores do not necessarily restrict themselves to the $1 price, they are united by their ability to provide low-price merchandise at fixed price points. To document this and other key characteristics of the dollar store, we regress prices – both for the good and per unit – on store type using the consumer panel for different aggregations of goods:

$$log(y_{kct}) = StoreType_{kct} + \alpha_{kct} + \epsilon_{kct} \tag{1}$$

$$log(y_{jct}) = StoreType_{jct} + \gamma_{jct} + \varepsilon_{jct} \tag{2}$$

where k is the upc/barcode, j is product module, c is the county, and t is month-year. y_{kct} and y_{jct} are our outcomes of interest. Our fixed effects specification allows us to observe differences across stores for the same product in the same location within the same month. As a result, our analysis is restricted to goods that are available across stores.

Prices across stores is the first outcome of interest, and Table 1 documents that dollar stores are able to offer significant price reductions. On average, dollar stores charge 11 percent lower prices for products with identical barcodes (as compared to grocery stores), as shown in column (1). This price reduction far outpaces discount stores (which includes Wal-mart), which are only 3 percent cheaper on average than grocery stores. These reductions are even bigger when we compare prices for the same product module, rather than barcode. As seen in column (2), the average product price at the module level is 47% lower at dollar stores relative to other retailers.

This price reduction is halved to 24% when the outcome is price per unit (i.e. an ounce of beef soup), as shown in column (3)[15]. That is, dollar stores offer package sizes within the same product module, a finding confirmed by the same regression with log size of the good as the outcome. As shown in Table A.1, dollar store goods are 24% smaller for identical product module units compared to other store types in the same county during the same month.

Beyond the sheer size of the price and size reductions, dollar stores also stand apart in how heterogeneous these effects are. Figures A.6 and A.7 show dollar stores' price per unit and size effects for each product group. While discount stores are quite uniform in their price and size effects, dollar stores vary greatly from product to product. Notably, almost all dollar store products are on average cheaper than those same goods at grocery stores. With the exception of milk, ice cream, and gum, dollar store shoppers receive discounts on every other product, even on a per unit basis.

Next, we turn to product variety across store types. First we collapse the data down to count the number of unique barcodes in a county and run the following regression:

$$log(N_{sct}) = a + \beta StoreType_{sct} + \alpha_{ct} + \varepsilon_{ct} \tag{3}$$

[15] There are no differences in size within the same barcode.

Table 1: Price Effects of Dollar Stores

Dependent Variables:	log(Price)		log(Price per Unit)
Model:	same bar code	same product	same product
Variables			
Dollar Store	-0.1106	-0.4655	-0.2384
	(0.0156)	(0.0096)	(0.0089)
Discount	-0.0311	0.0410	-0.0067
	(0.0017)	(0.0024)	(0.0022)
Club	0.4359	0.9212	0.1984
	(0.0197)	(0.0095)	(0.0060)
Convenience	-0.00006	-0.1500	0.0825
	(0.0091)	(0.0073)	(0.0089)
Drug	-0.0733	-0.1980	-0.0655
	(0.0032)	(0.0054)	(0.0034)
Fixed-effects			
county_upc_month	Yes		
county_product_month		Yes	Yes
Fit statistics			
Observations	479,718,710	479,718,710	479,718,710
R^2	0.92327	0.54447	0.77713
Within R^2	0.01631	0.11964	0.00636

Clustered (county_name) standard-errors in parentheses

Notes: Table reports coefficient from regressing log price per good and price per unit of good on a store type variable. We only use sales from dollar stores, discount stores, club stores, and drug/convenience stores. Grocery stores are used as the reference group. Data is based on consumer panel microdata for years 2008-2018. Column (1) reports coefficients with county by barcode by month-year fixed effects. Column (2) and (3) report coefficients with county by product module by month-year fixed effects. Standard errors are clustered at the county level.

where N_{sct} is the number of unique barcodes in county c at month-year t at store type s. We control for county by month-year fixed effects. We also examine number of product modules and number of product groups as the outcome variable.

Grocery stores by far have the most variety in every product aggregation level, as shown in Table 2 below. Grocery stores are the reference store type in the regression; compared to dollar stores, groceries have 20 times more barcodes, 10 times more product modules, and 5 times more product groups available. Dollar stores variety offerings are on par with that of Club stores, much less than discount stores, but far more than convenience and drug stores.

Dollar stores are not only offering less variety, but product offerings also differ. Figure 2 shows the top six product groups, as ranked by consumer expenditure share at each store. Grocery

Table 2: Variety Effects of Dollar Store

Dependent Variables:	log(No. of UPCs)	log(No. of Modules)	log(No. of Product Groups)
Model:	(1)	(2)	(3)
Variables			
Dollar Store	-2.871	-2.341	-1.643
	(0.0807)	(0.0590)	(0.0478)
Discount	-0.6223	-0.4018	-0.2050
	(0.0717)	(0.0483)	(0.0285)
Club	-2.649	-1.990	-1.308
	(0.0702)	(0.0564)	(0.0452)
Convenience	-4.696	-3.926	-2.910
	(0.0990)	(0.0715)	(0.0578)
Drug	-3.515	-3.059	-2.248
	(0.0427)	(0.0404)	(0.0489)
Fixed-effects			
county_name-state_name-month_year	Yes	Yes	Yes
Fit statistics			
Observations	1,528,436	1,528,436	1,528,436
R^2	0.83265	0.81585	0.76054
Within R^2	0.77464	0.76376	0.70110

Clustered (state_name) standard-errors in parentheses

Notes: Table reports coefficient from regressing log variety in a county-month-year on a store type variable. We only use sales from dollar stores, discount stores, club stores, and drug/convenience stores. Grocery stores are used as the reference group. Data is based on consumer panel microdata for years 2008-2018. Standard errors are clustered at the county level.

store consumers purchase large amounts of cheese, deli meats, and fresh produce, offerings typically not found at a dollar store. In fact, dollar stores look much more similar to discount stores.

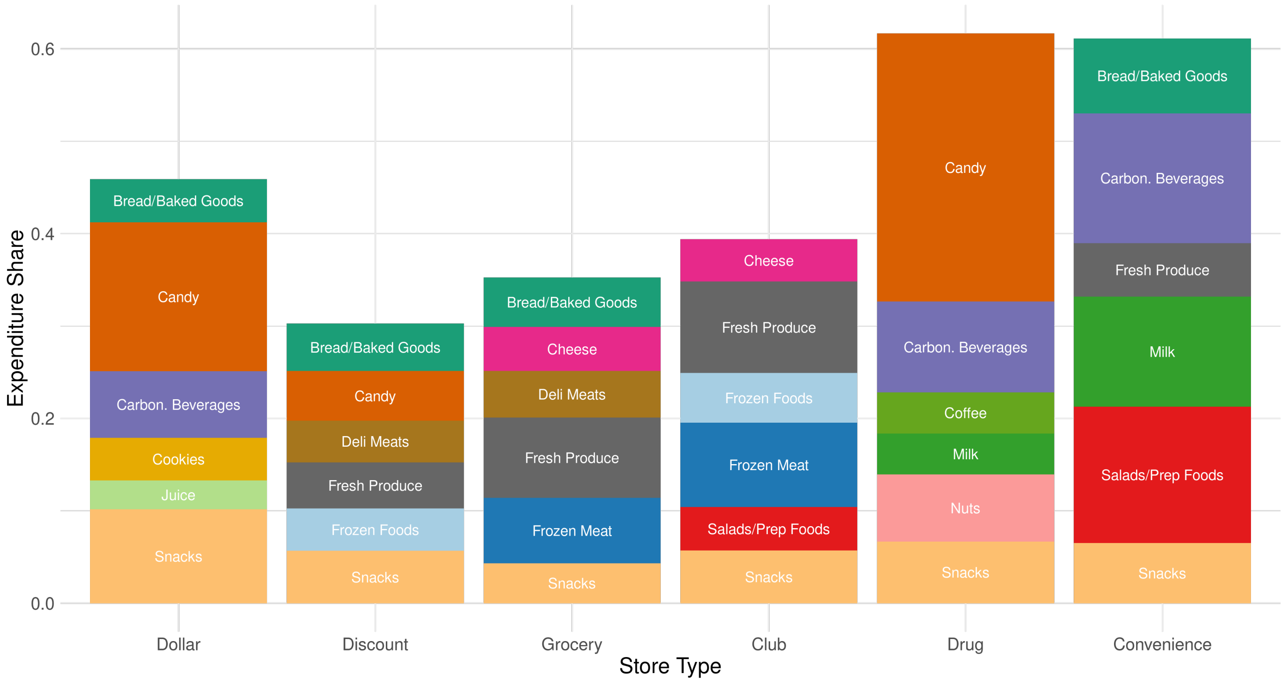
Taken altogether, a picture emerges of an average dollar store, which offers a limited selection of products, a limited variety of brands within those products, likely at smaller sizes than available at other retailers, but all in all at a much lower price than anywhere else.

4 Empirical Analysis of the Impact of Dollar Store Entry

Our goal is to understand the effect of dollar stores on consumers and preexisting retailers. In theory, the dollar store can indirectly impact the consumer by putting competitive pressures on local rivals to change prices or exit the market.

We leverage an event study design to examine supply-side changes in response to the first dollar store entry in a zipcode. To investigate the grocery count in a zipcode, we run the following

Figure 2: Expenditure Share of 6 Most Popular Products by Store Type



Notes: Figure reports consumer expenditure shares of the top 6 product groups for each store type. Data is based on consumer panel microdata for years 2008-2018.

regression:

$$Y_{zt} = \Sigma_{k=-T_1}^{-2} \beta_k \times D_{zk} + \Sigma_{k=0}^{T_2} \beta_k \times D_{zk} + \gamma \boldsymbol{X}_{zt} + \nu_z + \phi_t + \varepsilon_{zt} \quad (4)$$

Let Y_{zt} denote grocery count in zipcode z in year t, D_{zk} is the years before or after entry of the first dollar store in a zipcode, $\boldsymbol{X}_{zt}$ are lagged zipcode level demographic controls, ν_z are zipcode fixed effects, and ϕ_t are year fixed effects.[16] Our parameter of interest is β_k, the amount by which the average zipcode experiences a change in number of grocery stores upon entry of the first dollar store into their zip code. Standard errors are robust and clustered by zip code.

We run an analogous household-level regression at the quarterly level, q, to investigate visits to other retailers and changes in prices:

$$Y_{iqt} = \Sigma_{k=-T_1}^{-2} \delta_k \times D_{ik} + \Sigma_{k=0}^{T_2} \delta_k \times D_{ik} + \sigma_i + \tau_{qt} + \epsilon_{iqt} \quad (5)$$

Here, conditioning on σ_i allows us to look at variation within households, rather than variation across households, and replaces demographic controls. Standard errors are robust and clustered by zip code.

The control group here is thus the not-yet-treated group. For estimation, we use heterogeneity-robust estimators developed by Callaway and Sant'Anna (2021). This method alleviates concerns over bad control groups, as discussed by the recent literature on staggered roll-out and two-way fixed effect (TWFE) designs (Baker et al., 2022).

[16] Controls include average household income, proportion of households married, average household size, average age, proportion white, proportion black, and average working hours.

To ensure the cleanest identification, the event is defined as the first dollar store entry in the zip code. That is, the control group comprises the not-yet treated, which is identical for each event. Had the event been defined as higher dollar store entry (e.g. 2nd or 3rd dollar store entry), household substitution between dollar stores within the same zip would likely have contaminated the resulting dollar store effect. Although the first dollar store entry is not without contamination – household could have shopped at a dollar store outside of their zipcode in the pre-period – the first entry provides the cleanest identification possible given the data available. One downside of this event choice is that we do not speak to store competition or household response in the presence of many dollar stores.[17]

The assumptions required for the event study design are no anticipation and common trends. While dollar store entry maybe be announced a quarter or so in advance, it is likely that households would not adjust their consumption until the dollar store actually enters. Furthermore, anticipation would likely induce a change in outcomes before entry, but pre-trends are flat. Our identifying assumption is that households in different zip codes that receive dollar stores in different times but will eventually receive a dollar store would have followed the same pattern absent dollar store entry.

A common concern with the event study strategy is dollar store entry is related to other features of the local retail environment that would affect household consumption patterns. However, if dollar stores respond to changes in local demand conditions, household consumption patterns would likely change even before the dollar store enters. To test for changing patterns before dollar store entry, we estimate the treatment effect in the eight quarters leading up to the entry of a dollar store. We find a precisely estimated flat pre-trend before dollar store entry for all outcomes, and a significant trend break at the time of the entry.[18]

4.1 Results - On average, no evidence of grocery store exit

Does the first dollar entry crowd out the grocery store? This concern stems from media attention on dollar store impacts and evidence of crowd-out in the big box store literature and the media[19]. Dollar stores generally do not stock perishable foods, and thus lack inventory in high nutrition fresh foods. Thus, if dollar stores crowd out grocery stores, this would not only decrease the number of available retailers, but would also limit access to healthy food items. In addition, our results in the next section show that dollar store entry induces substitution away from grocery stores, and thus we should be most concerned that grocery stores are edged out when dollar stores enter.

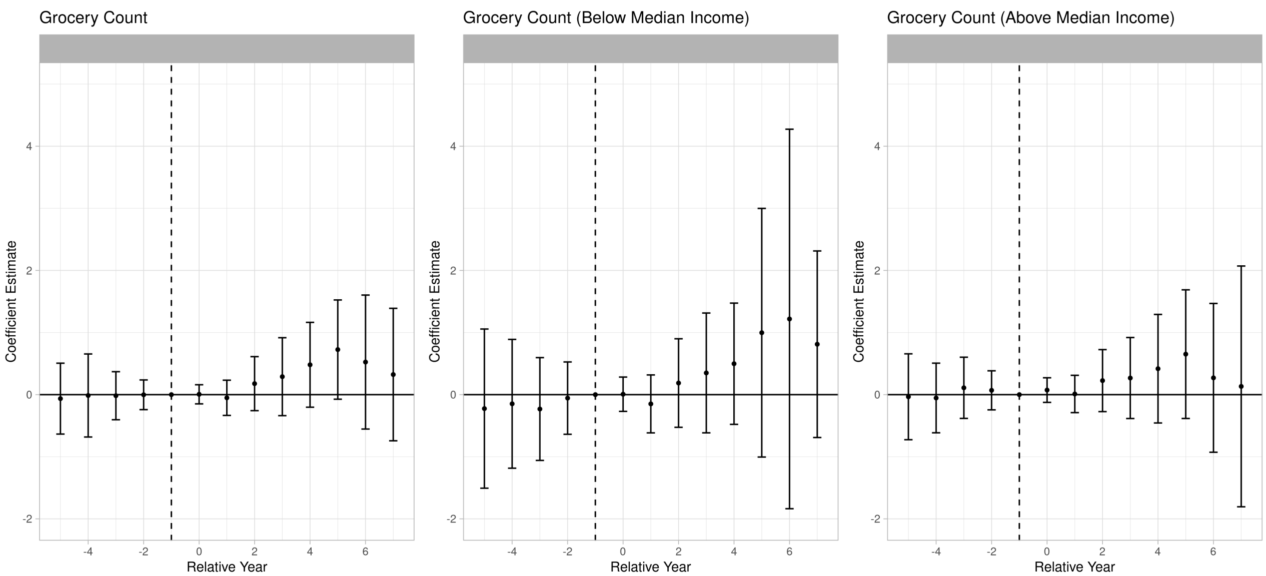
[17]Policymakers who have banned dollar store entry tend to operate in areas with at least fifty dollar stores. We do not claim to speak to the normative implications of this many dollar stores in an area.

[18]Additionally, we show the event study results with 5-year average zipcode income (using American Community Survey data) as the dependent variable, shown in Figure A.9. We do see some general decline in average income in a zipcode around the time when dollar stores enter. However, note that our event study includes household fixed effects, which would control for this effect.

[19]For example, this CBS News Article details this possible concern and other common concerns regarding dollar stores.

Our event study analysis from Equation 4 provides no evidence of this crowd out effect for the first dollar store. As shown in Figure 3, there is no change in the number of grocery stores after the first dollar entry[20]. This pattern holds across neighborhoods with different socioeconomic characteristics, as shown by the second and third panels. Thus, on average, there is no grocery exit after the first dollar entry.

Figure 3: Effect of First Dollar Store Entry on Grocery Count



Notes: This figure reports event study estimates with 95% confidence intervals from Equation 4, using 2008-2018 SNAP and ZBP data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). The figure reports the grocery count as the outcome variable. Errors are clustered at the zip code level.

This result is distinct but consistent with preexisting literature: for example, using ordinary least squares with controls, Caoui et al. (2022) find that the first dollar store entry causes a 7% decline in independent grocery stores within two miles of the population center of the census tract, a small value which is consistent with an overall change with the number of grocery stores.

4.2 Results - No evidence of price competition

Price competition is a potentially important aspect of the consumer and local retailer response to the dollar store. Intuitively, the price of the consumer bundle may change due to competition with preexisting retailers or because household choice set expands to include new and cheaper goods. On the supply side, prices may change as a result of increased competitive pressure from an additional store or competition with the dollar store's lower prices (as shown in Table 1 and

[20] As robustness, we run the same event study without demographic controls in Figure A.18, and there is still no grocery store exit. Counterintuitively, there appears to be a slight increase in grocery store count after a dollar store enters, although this not statistically significant at the 95% level, and is driven entirely by above-median zip codes

Figure A.6, dollar store prices are lower on almost every product group). On the demand side, exposure to cheaper goods might induce households to switch to the cheaper dollar store option (within UPC/barcodes) or might induce to switch to a cheaper variety (across UPC/barcodes). However, if the price of the household bundle does not change in response to the dollar store entry, then there is not a big competitive price response.

To understand the effect of the first dollar store on price of the household bundle, we therefore construct a relative price index (RPI) for each household, following Aguiar and Hurst (2007) and Leung and Li (2021). Intuitively, the relative price index is the household's expenditure relative to the household's counterfactual expenditure wherein the household purchases the same goods but at an average state price:

$$RPI_{ist} = \frac{\sum_{j \in J_{it}} p_{jist} q_{jist}}{\sum_{j \in J_{it}} \bar{p}_{jst} q_{jist}} \tag{6}$$

p_{jit} and q_{jit} are the price and quantity for product module j for household i at quarter-year t in state s. The numerator in this expression is thus the total expenditures for household i in quarter-year t. The denominator in this expression is the total expenditure if prices paid are replaced by the state average price in that product. We will refer to the denominator as the "counterfactual expenditure", which is constructed by calculating a reference price, $\bar{p}_{jt}$ for each region s the household is located in.

$$\bar{p}_{jst} = \sum_{i \in I, d \in t} p_{jist} \left(\frac{q_{jist}}{\bar{q}_{jst}} \right) \tag{7}$$

where $\bar{q}_{jst} = \sum_{i \in I, d \in t} q_{jist}$ is used for weighting the price by the quantity purchased of the product.

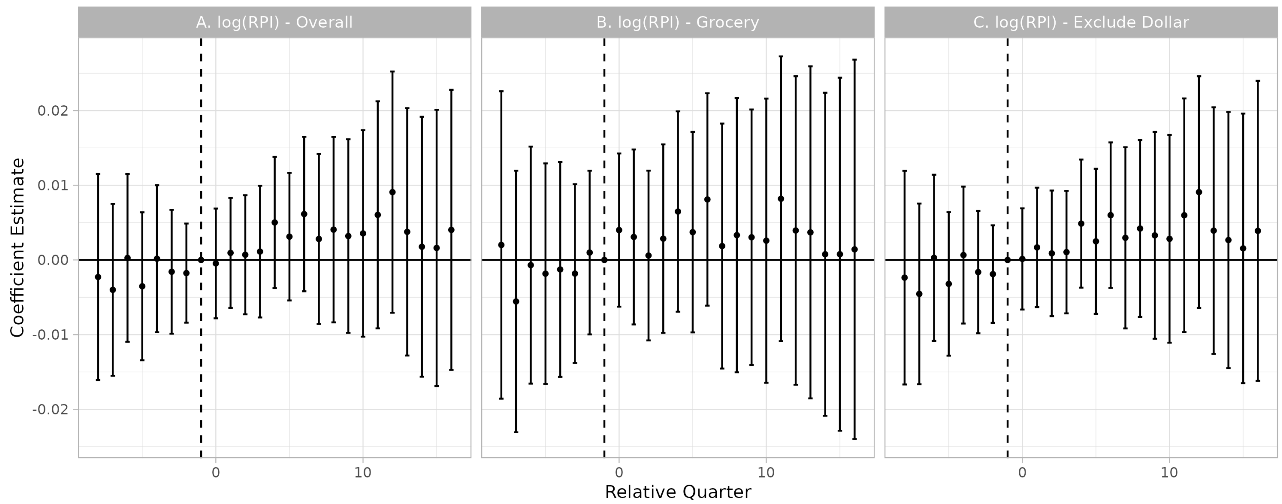
The RPI essentially compares a household's true expenditure to a "fixed" counterfactual expenditure. A decrease in RPI after a dollar store enters would imply that entry induces a lower priced bundle for households, as compared to the same bundle of goods in other parts of the state. This could be either due to dollar stores offering cheaper prices, and/or a competitive response from other stores as a result of dollar store entry.

In Figure 4, we find a surprising and precise null result of dollar store entry on the RPI. That is, following dollar store entry, households pay the same amount for a representative good as households in the same state, and the confidence interval range between ± .02 percentage points. We chose the state average price because a dollar store is quite small, and using the state average price ensures that the reference price we use is not reacting to the entry event. Note that a major difference between this result and the significant price differences we observed in Table 1 is that the price difference regressions focused solely on products sold across stores, while the RPI analysis pools all goods purchased both before and after dollar store entry. This lack of price reaction is consistent with Arcidiacono et al. (2019), who also finds a null incumbent response from Supercenter entry.[21] Interpreting the results, this lack in price change

[21] Arcidiacono et al. (2019) argue that the explanation for this is stores routinely employs cost-plus pricing, or markup pricing, where a fixed percentage is added on top of the unit cost of a product.

in the consumer bundle implies that the reduction in household total expenditures comes from a switch in varieties/UPCs or a decrease in quantities within a variety, not a change in price. We also note that the null RPI response reflects a switch from a more expensive variety to a cheaper variety, but does not reflect the additional price cut from switching to the dollar store (from Table 1, on average, dollar store prices are 10% cheaper than grocery prices for the UPC and 25% cheaper for the product module).

Figure 4: Effect of First Dollar Store Entry on Prices - State Average Reference Price



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). The dependent variable, the Relative Price Index (RPI), is shown in Panel A and is defined in Equation 6. Panel B and C repeat the analysis, with data restricted to purchases at the grocery store and non-dollar store, respectively. Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

We further test whether there is a competitive price response by restricting the sample of expenditures a household makes to only grocery purchases (Figure 4 Panel B) and non-dollar store purchases (Figure 4 Panel C). Both these sets of analyses show a null RPI response. The null result in grocery prices shows us that on average, the household pays the same prices relative to the state price purchases at the grocery store before and after dollar store entry. Generalizing this to all non-dollar stores, the prices paid by consumers do not change, on average, after entry.

Combining the lack of price movement with the lack of grocery store exit and lack of trade-off in shopping trips between store types, we conclude that in response to the first dollar store, the supply-side of the retail market remains relatively fixed, except for the added choices the dollar store provides.

4.3 Results - On average, households still visit grocery and discount store types at the same rate

We might expect that dollar store entry could change household choice sets, or induce reactions by incumbent retailers, if households replaced some trips to other store formats with dollar store trips. To test this, we estimate the effect of dollar store entry on the number of trips to each store type using the specification in Equation 5. Here, the analysis focuses on the effect of entry on grocery stores and discount stores for two reasons. First, from the policy perspective, substituting away from stores that consistently carry healthy food options (i.e. grocery and discount stores) is a major policy concern. Second, as we will see in Section 5, dollar store entry accompanies a drop in expenditures at grocery and discount stores, but not other store types.

As shown in Figure 5, on average, households take more trips to the dollar store without significantly changing the number of trips to grocery stores and discount stores. Specifically, dollar store trips increase by 5 percentage points each quarter on average, and the effect is significant and persists at least four years after the first dollar store entry. Meanwhile, other trips (overall, to the grocery store, to the discount store), decrease slightly in the third and fourth quarter following dollar store entry but revert back afterwards. The total number of trips do not change after dollar store entry, and the effect is a precise null (the bounds on our estimates are between $\pm$ .04 percentage points each quarter). As a result, trips to other store types decrease, but this decrease is distributed over several store types in such a way that the number of trips to each store types does not change significantly. That is, the lack of significant change in number of trips to other store types provides evidence that on average, households have the opportunity to purchase the same products after the first dollar store entry. Along with a lack of grocery exit, the lack of trade-off in shopping trips between stores types implies that the household choice set is (weakly) increasing following the first dollar store entry.

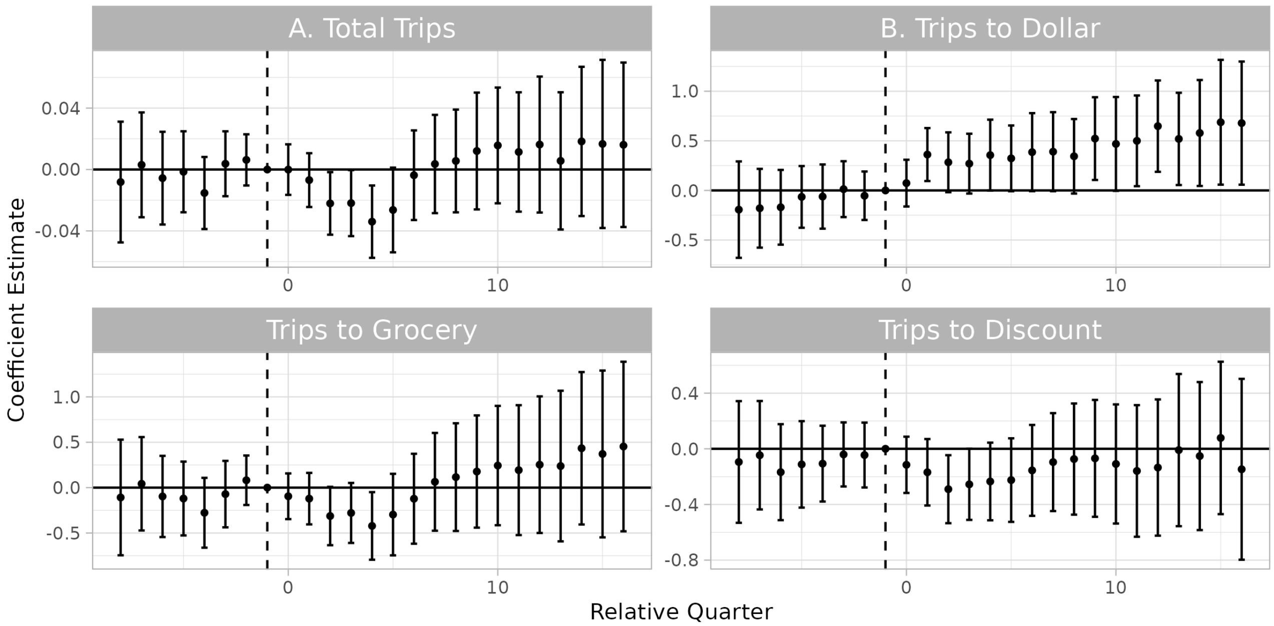
In this section, we established the effect of the first dollar store on the local retail market and incumbent stores. A main policy concern is that dollar store entry could lead to supply-side changes with ambiguous welfare impacts. We presented three pieces of evidence that the consumer choice set remains the same after the entry of the first dollar set: (1) grocery count remains the same, (2) households still visit other retailers at the same rate, and (3) no evidence of change in prices.

5 Effect of Dollar Store Entry on Household Consumption

We examine how the dollar store affects household consumption. Since the first dollar store does not induce a significant supply-side response, the first dollar store entry captures the direct effect of the dollar store format on consumers.

First, we document that the dollar store leads to a decrease in total food expenditures. As shown in Figure 6, total food expenditures are flat in the lead-up to the dollar store entry

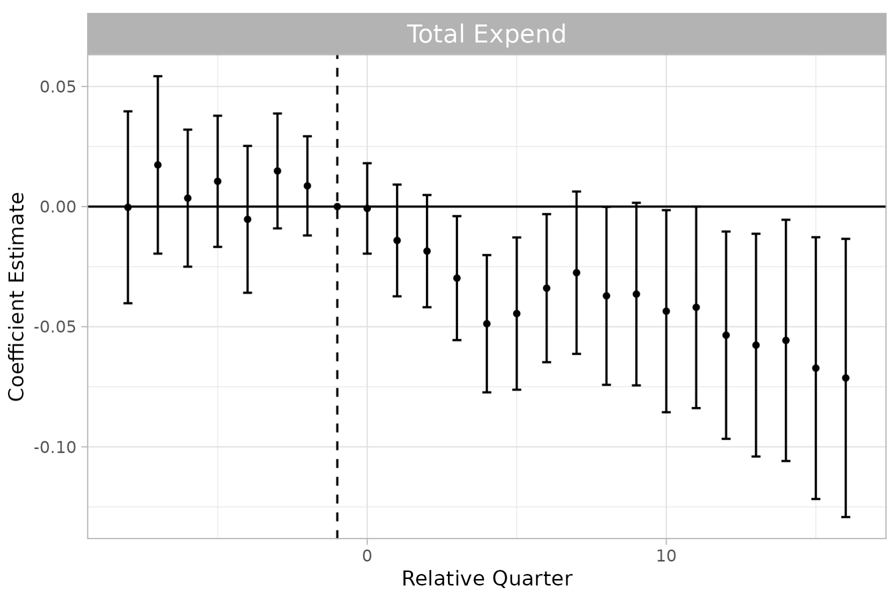
Figure 5: Effect of First Dollar Store Entry on Households' Log Number of Trips



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

event, start falling as soon as the entry occurs, and stabilize at a negative and significant 5 percentage points. This drop continues even after 16 quarters, demonstrating a persistent effect.

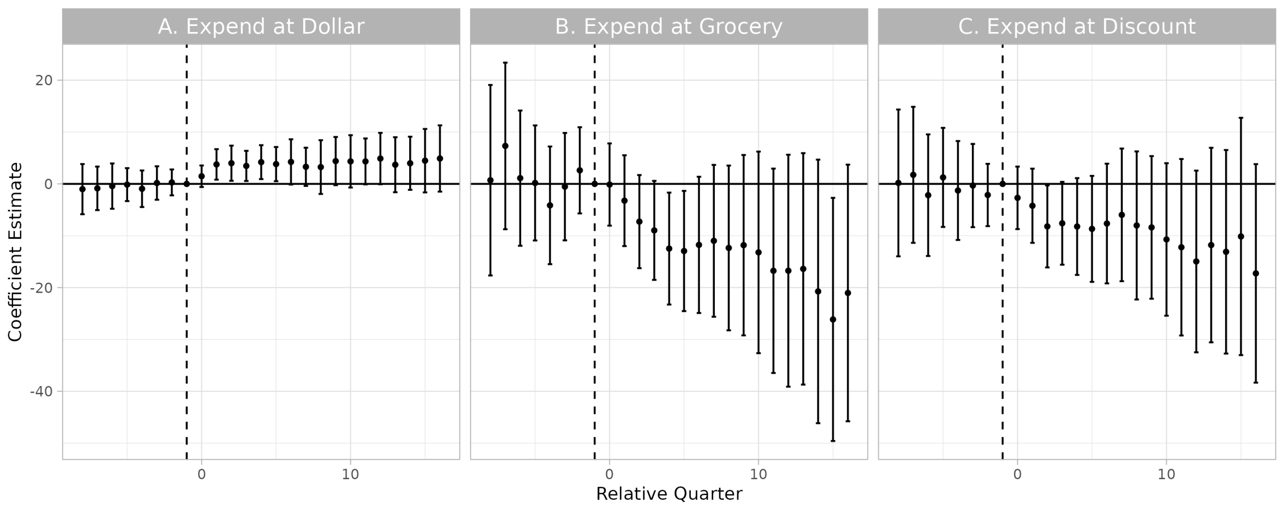
Figure 6: Effect of First Dollar Store Entry on Households' Total Log Food Expenditure



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). The figure reports log of total expenditures of food. Observations are not weighted for national representativeness. Errors are clustered at the zip code level.

This overall decline stems from substitution from grocery stores and discount stores to dollar stores, as seen in Figure 7. The decline in total expenditures is driven by decreased expenditures at grocery and discount stores, which outweighs the increased expenditures at dollar stores. Figure A.10 shows the expenditure response from all six store types of interest. Superstore/Club, convenience, and drugstores do not exhibit a response from entry, and we thus focus our discussion on grocery and discount stores.

Figure 7: Effect of First Dollar Store Entry on Households' Expenditure at Various Store Types



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

We turn our focus to understanding the features of the dollar store format that could explain this drop in expenditure. For household i, total expenditure at any given time is:

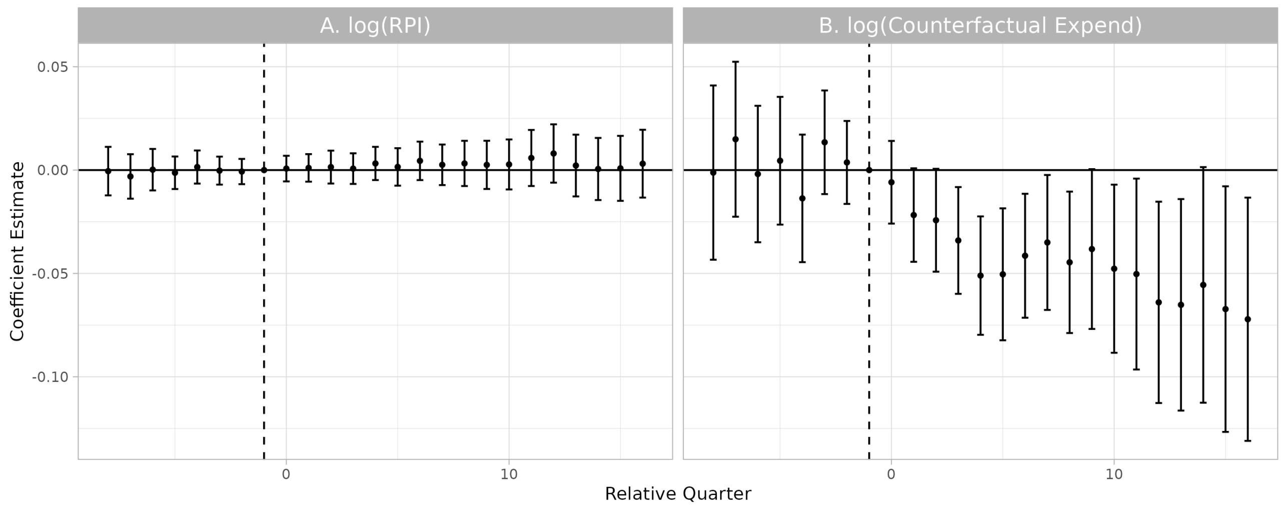
$$E^i = \sum_{j \in J} p_{ij} q_{ij}$$

Mechanically, three distinct components could change such that total expenditures would drop after a dollar store entry: (1) net drop in price of good j, (2) net drop in quantity of good j, and (3) change in product choices. In the following sections we explore if each of these three mechanisms contribute to the drop in expenditures. In section 6 we provide a theoretical framework to quantify how these changes translate into welfare. Previewing our results, we find no evidence that the households are buying the same consumption bundle at lower prices. Instead, the expenditure drop is driven by quantity changes: consumers are buying fewer amounts and shifting towards lower priced varieties.

5.1 The price of the household basket of goods is unchanged

We have already shown that decreases in total expenditure cannot be explained by households enjoying lower prices for the same products they previously were consuming. This was demonstrated by the null movement in RPI shown in Figure 4. We repeat this exercise, but with the reference price set at the county level so that we are comparing household expenditure to prices offered for the same good, but at the county average price. If households were purchasing the same goods in the post period as in the pre period, and dollar stores were offering cheaper prices, we should observe a drop in the RPI post-entry. The null result we see in Figure 8 provides evidence that this is not the case.

Figure 8: Dollar Store Entry on Prices - County Average Reference Price



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). The dependent variable, the Relative Price Index (RPI), on the left side panel is defined in Equation 6. The dependent variable on the right side panel is the counterfactual expenditure, where the reference price is defined in Equation 7. Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

To better understand this result, we also show the counterfactual expenditure (at the county level), which is the denominator from RPI Equation 6. The counterfactual expenditure is computed as the the household expenditure where prices of each good are replaced by state average prices. As shown in Figure 8, the counterfactual expenditure declines post-entry. In fact, this decline in the counterfactual expenditure mirrors the decline in total food expenditure shown in Figure 6.[22] Assuming that the reference price is unaffected by dollar store entry (said otherwise, assuming that dollar store entry does not affect average food prices at the county or state level), decreases in counterfactual expenditure suggest that decreases in expenditures are driven by quantity changes, variety changes, or both.

5.2 Households reduce the quantity of goods purchased

Next, we analyze whether the first dollar store causes households to purchase less quantity on net. Quantities could decrease if prior to the dollar store entry, the household lacked smaller-sized options and thus purchased more than the optimal amount of goods. In this case, the dollar stores' small-sized products would allow the household to re-optimize and thus reduce the quantity purchased. To measure quantity in observable units, we compare ounces consumed

[22] The figure is almost the same when the reference price is at the state level.

for different product groups before and after the first dollar store entry.[23] To give a complete picture of food shopping behavior, we measure quantity at the department level, the highest level of aggregation for Nielsen.

Figure 9 shows the effect of dollar store entry on households' quantity purchased using the same event study analysis from Equation 5. No departments show increases in quantities following the entry of the first dollar store. Three departments – fresh produce, deli, and packaged meats – show no change in quantity, although the error bars are large (± 10 percentage points) and so the result is noisy. The remaining three departments – dairy, dry grocery, and frozen food – show a 4% reduction in the average quantity consumed. While the drop in quantity is temporary for dairy and frozen foods, the reduction is persistent for dry groceries.

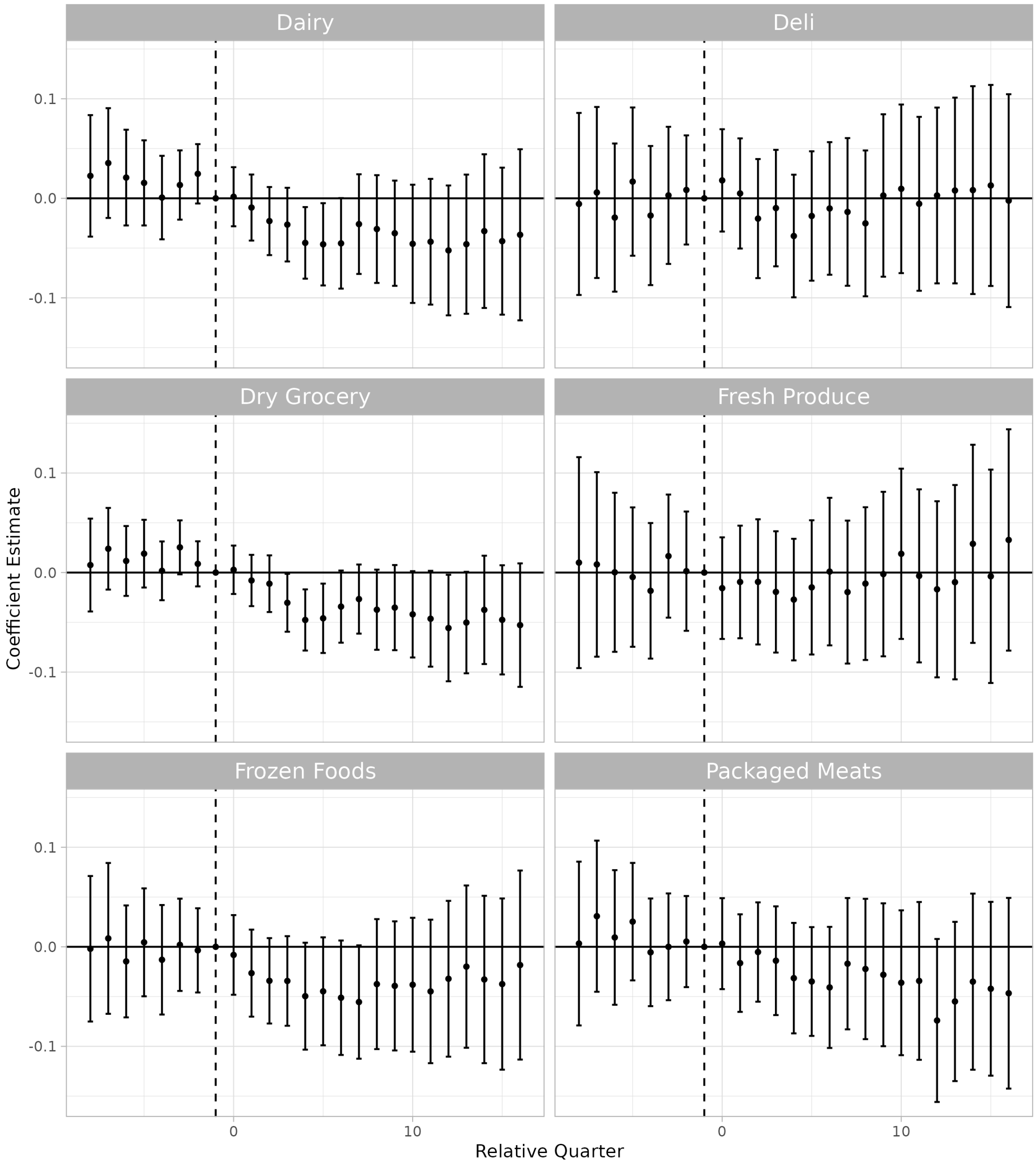
Households most reduce quantities for dry goods, frozen foods, and dairy (as shown in Figure 9), the three most popular departments at the dollar store by expenditure share (as shown in Figure A.13). Interestingly, while households decrease their overall expenditure in dry goods (by about 10 dollars per quarter), the household expenditure share shifts towards dry goods, as shown in Figure A.12. These results are all consistent with the hypothesis that before the dollar store, households were not optimizing on the size of the goods they were purchasing. The dollar stores' small-sized products allow household to reoptimize and thus reduce the quantity they purchase.

Important for policymakers is the result that households do not significantly change the quantity of fresh produce, speaking to concerns that the dollar store format causes unhealthy eating by inducing a substitution from broadly healthier categories (like fresh produce) to broadly less healthy categories. To investigate further, we also show results of the quantity analysis featuring the ten most popular product groups for dollar, grocery, and discount stores (the most popular categories are shown in Figure 2). Figure A.17 show null changes in quantity for all the selected product groups. These results hold both for more processed items that the dollar store specializes in (shown in purple), as well as for less processed items that are not commonly found at the dollar store (shown in green). These nulls suggests that if the first dollar store impacts consumer health, then the effect due to substitution towards away these broad product categories is small.[24]

[23]The majority of product sizes in the HMS dataset are measured in ounces. The liquid measurements are in "fluid ounces" and the weight measurements are in "ounces". We convert quarts and pounds into fluid ounces and ounces. When possible, we convert "counts" into ounces using an average measure – for example, one egg weighs 1.7 ounces. Data measured in counts that cannot be converted into a weight is excluded from the measurement.

[24]Our "health" result is only a very crude approximation for health. Allcott et al. (2019) uses a Healthy Eating Index (HEI), a multi-dimensional measures that calculates the overall nutritional intake of each food. Further analyses on the nutritional content lies outside the scope of this study. In addition, we do not investigate within product-group changes in health.

Figure 9: Effect of First Dollar Store Entry on Log Ounces of Each Department

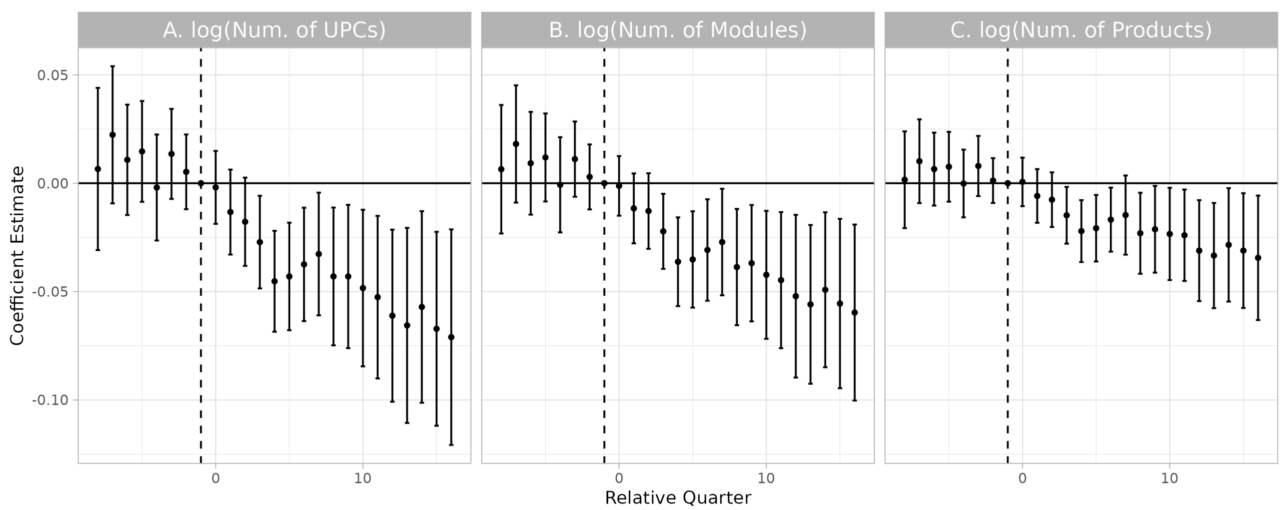


Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

5.3 Expenditure Drop Stems from Changes in Product Choices

Finally, we report the effect of the first dollar store on household variety in Figure 10. We investigate three different measures of aggregation: number of distinct UPCs purchased in one quarter, number of distinct modules purchased in one quarter, and number of distinct product groups purchased in one quarter. For all three levels of aggregations, pre-trends are flat in the lead up to dollar store entry, variety declines following dollar store entry, and the effect is persistent. From most disaggregated to most aggregated, UPC/barcode variety captures depth of variety, whereas the number of products groups captures breath of variety. We find persistent and significant declines of unique varieties at all levels of aggregations: UPC/barcodes decline by 6%, product modules decline by 5%, and product groups decline by 1%.[25] That is, we observe that the dollar store shifts consumption towards lower-priced goods, at the expense of variety. We also observe that dollar stores are the store type with the least variety. We quantify how households value price, variety and other product characteristics in the demand estimation.

Figure 10: Effect of First Dollar Store Entry on Consumption Variety



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). The dependent variables here are different definitions for consumption bundle variety for the household at the quarter-year level. Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

[25] For context, the average household in the reference period (k = 1) consumes 111 distinct UPCs, 64 distinct product modules, and 33 distinct product groups.

6 Theoretical framework

We use a model to understand how consumers value aspects of the dollar store bundle, to rationalize our empirical findings, to quantify the welfare effects of the dollar store, and to address policy questions about the dollar store's effect on the consumer. We estimate demand to determine which of the dollar store characteristics (documented in Section 3) drive the change in the household consumption bundle (documented in Section 5). Then, these estimates are used to assess the price and quantity/variety trade-off observed in the empirical analysis. Then, the demand estimates are used to rationalize the reduction in expenditure, quantity, and variety. More generally, the demand estimates shed light on the demand-side drivers of dollar store proliferation.

We study whether the dollar stores' provision of low cost product groups pushes households to trade-off fresh produce at the grocery store (a plausibly healthy product group) for product categories at the dollar store (some of which nutritionists have argued are ultra-processed and unhealthy). In the context of the empirical analysis, we test whether dollar store entry decreases ounces of fresh produce consumed or increases ounces of arguably unhealthy product groups. In the context of the demand estimation, we test whether broad product categories at different store types are close substitutes. For this reason, we model broad levels of product aggregation: product group at a store type. To accommodate the fact that different stores have different varieties, sizes, and prices within product groups, we treat the product group at each store as an amalgamation of the bar codes available in that product group in that store. Then, the representative household might face a choice between a more expensive product group at the grocery store with bigger sizes and more variety and a cheaper product group at the dollar store with smaller sizes and less variety.

In addition, the model needs to be consistent with the reduced form results. Specifically, the model needs to accommodate reductions in expenditure, counterfactual expenditure, quantity, and variety. In our model, these observations constitute a switch from a higher-priced product comprised of more unique varieties and larger sizes to a lower-priced product comprised of less quantity and fewer varieties. The observed reduction in quantity (amount) and the lack of relative price change is consistent with the introduction of new goods that have the same relative price and that ultimately weigh less than the previous product.

We estimate demand using a discrete-choice, nested logit model.[26] In this model, a household first chooses a store type to shop at – grocery, dollar, club, convenience, discount, or drug, and then chooses which product group to buy within that store (a product group is comprised of bar codes within that product group). Then, household preferences are based on product characteristics (such as price, variety and size), a product group specific preference, as well as a household idiosyncratic shock, where the nests are the store types. Then, the household utility function is represented as

[26]This nested logit model can be represented either as a household choosing one product group within a store, or as a representative household that purchases shares of each product group, for details see Verboven (1996).

$$u_{ij} = -\underbrace{\alpha \log p_j}_{price} + \underbrace{\beta v_j}_{variety} + \beta_2 v_j^2 + \underbrace{\eta m_j}_{product\ size} + \eta_2 m_j^2 + \underbrace{\psi_{g(j)}}_{product\ group} + \xi_j + \epsilon_{ij}(1-\lambda) \quad (8)$$

where each household has utility over product j, product group g at store s, that depends on the price p_j, product characteristics observable to the researcher, v_j, p_j, m_j, and characteristics unobservable to the researchers which are potentially endogenous, ξ_j and idiosyncratic shock ϵ_{ij}. The nest parameter, λ, indicates the degree of substitution between products within nest.[27]

The "product" is an amalgamation of bar codes within that product group and store. To aggregate from bar code to product group, we assume that households' preferences follow a Stone price index, as in Atkin et al. (2018), or that the household consumes all varieties within the product group, and pays an expenditure-weighted sum of log prices for these varieties that comprise the single product group:

$$\log p_j = \sum_{b \in j} \phi_b \log \tilde{p}_b \quad (9)$$

where p_j is the price of product j, which is comprised of bar codes b, ϕ_b is the household's expenditure on bar code b divided by the household's total expenditures on product j (product group g at store s), and $\tilde{p}_b$ is the price paid for bar code b. Then, the price for product group g at store s is the store-level fixed effect. As a result, the price we measure is a relative price in the market, and is measured in log dollars.

Then, following this product definition, we compute variety v_j as the average number of unique bar codes for product group g at store s and size m_j as the average size of the bar codes within a product group within a store. Since dollar stores offer small sizes and fewer varieties, we include second order terms to account for decreasing marginal returns and to test whether dollar stores are valued on this margin.

Since there is lack of supply-side response, the model focuses entirely on demand and on the direct effects of the first dollar store entry. Had there been a competitive response from grocery stores or other store types, we would have decomposed welfare into the direct effect from the dollar store and the indirect effect from competing retailer response (for an example of this, see Atkin et al. (2018)).

[27] Our model is in line with the existing literature but tailored to the empirical results. Allcott et al. (2019) model grocery demand as a choice of product groups using Cobb Douglass utility and measure consumption in calories. Handbury (2021) model product group choice using a combination of Cobb Douglass and log-logit preferences. Atkin et al. (2018) model households as first choosing a product group and then choosing a store within product group, with Cobb Douglass preferences over product groups and CES preferences over stores within product groups. Atkin et al. (2018) use a Stone price index to aggregate from bar code to store product group. We follow Atkin et al. (2018) to aggregate from bar codes to product groups within stores, but use a nested logit model to capture the reduction in total food expenditure following dollar store entry.

7 Estimation and the Value of the Dollar Store

The model is taken to the data. Intuitively, we compare relative prices and shares of product groups in different stores across different markets in order to identify demand parameters. We define the outside good as the set of weights for the products that are not in the top twenty product groups. Then, under the model, the quantity share of good j (product group g at store s) is given by

$$\log\frac{\pi_{jt}}{\pi_{0t}} = -\frac{\alpha}{1-\lambda}\log p_{jt} + \frac{\beta}{1-\lambda}v_j + \frac{\beta_2}{1-\lambda}v_j^2 + \frac{\eta}{1-\lambda}m_j + \frac{\eta_2}{1-\lambda}m_j^2 + \frac{\psi_{g(j)}}{1-\lambda} + (1-\lambda)\log\pi_{jt|s} + u_{jt} \tag{10}$$

where π_{jt} is the share of good j in market t, $\pi_{jt|s}$ is the share of good j in store type s in market t, and π_{0t} is the share of the outside good; $\log p_{jt}$ is a relative price of product group in its store relative to the prices of the same product group in other stores in that market, v_j and m_j are product characteristics and $\psi_{g(j)}$ are product group fixed effects.

Quantity shares and prices are determined simultaneously in equilibrium, and, in particular, demand shocks are likely correlated both with prices and shares. To overcome this endogeneity, we employ two instruments: the average price of the same good in other markets for the same retailer (following Hausman et al. (1994)), and we also introduce what we call the retail closing instrument (discussed in detail in the next section).

The average retailer price exploits the idea that local demand shocks are likely uncorrelated with prices in different markets. Intuitively, local pricing decisions can depend on both supply and demand factors, and the average price in different markets captures the supply component without capturing the idiosyncratic demand in a market. Furthermore, since prices for many retailers are set at the national level, it is even more likely that prices do not reflect local demand shocks (DellaVigna and Gentzkow (2019)). Even when retailers price uniformly, prices vary across markets as we employ a relative price index across firms in the same market, and different markets are composed of different stores.

To test for weak instruments, we regress shares (log shares of each good divided by the outside good) on log prices instrumenting with the average retail price in other cities and the retail closing instrument and controlling for the characteristics specified by our demand model. The first stage with F-test is shown in Table A.6.

Demand parameters are identified from variation in prices, shares, and characteristics. For estimation, we follow the approach from Berry et al. (1995) and Conlon and Gortmaker (2020).

7.1 Retail closing instrument

We introduce a "retail closing instrument" which exploits a plausibly exogenous shock that lowers the cost for a retailer to enter a zip code.[28] Throughout the 2000s and 2010s, technology shocks and other market forces caused a wave of bankruptcies that shut down several major brick and mortar retail chains (chains that are unrelated to food, such as Blockbuster).[29] These store closures created opportunities for other retailers to enter.

Since these closures occurred abruptly and bankruptcy occurred at the national level, we assume that the timing of these closures were unrelated to local demand for food. We thus leverage bankruptcies of retail chains that occurred during our study sample, and chose chains that were previously operating at a national scale, listed in Table A.2. These bankruptcies occurred at different points in time, as illustrated in Figure A.19, and local markets had heterogeneous exposure to each bankruptcy. The identifying assumption here is that retail bankruptcies only affect local food markets by creating a vacant storefront where another (potentially food-providing) store can locate. We then define the retail closing instrument as a binary variable equal to one if the county had a store that went through bankruptcy in the year(s) prior, and zero otherwise. We report the first stage with F-statistic for the one-year-lag bankruptcy instrument in Table A.6.

We illustrate the variation we are exploiting with an exercise focused exclusively on dollar store entry. We regress the first dollar store entry on a modified versions of the retail closing instrument described in the previous paragraph. In practice, it can take several years after bankruptcy to shut down stores, and, it can take several years to open a new store in a vacant location. We thus include further lags of the bankruptcy instrument. We additionally control for year fixed effects, as shown in the equation below:

$$D_{it} = \alpha + \beta_k Z_{ik} + \lambda_t + \epsilon_{it} \tag{11}$$

In this regression, D_{it} is an indicator variable for whether the first dollar store has entered zip code i by year t, and Z_{ik} indicates the presence of a bankrupt retailer in the zip code i k years before that the first dollar store entry to the zip code. We report results in Table A.7, which show that dollar store entry is positively correlated with retail closings.

7.2 Disentangling the relative importance of store characteristics

We report the demand estimates in Table 3. We model price, size, and variety directly in order to distinguish which features of the dollar stores are preferable to consumers.[30] We find that households prefer lower prices, larger sizes, and more variety, and observe slight decreasing marginal returns in variety and size. Intuitively, lower-income households are the most price

[28]We thank Fern Ramoutar for conceptualizing the use of retail bankruptcies as an instrument.

[29]For anecdotal evidence of what is driving the "retail apocalypse", see the New York Times, Washington Post, and Wall Street Journal.

[30]In addition, we consider the value of convenience in Appendix A.3.

sensitive and higher-income households are the least price sensitive. Preferences over variety and size are similar across income groups. Then, the relative strength and magnitude of these estimates clarify which aspects of the dollar store bundle (documented in the descriptive statistics) drive changes in household consumption (documented using the event study framework). Compared to other store types that feature larger sizes and more variety but higher prices, the estimates suggest that the only dollar store attribute that provides household utility are its low prices.

Table 3: Demand Estimates

Variable	Income Rank 1	Rank 2	Rank 3	Rank 4
price	-0.8515	-0.7891	-0.6994	-0.5141
	(0.01433)	(0.01107)	(0.009681)	(0.008042)
num varieties	0.04367	0.03623	0.03327	0.03128
	(0.0008513)	(0.0006342)	(0.0004869)	(0.0005691)
avg. size	0.005815	0.004137	0.003529	0.003277
	(0.0001937)	(0.0001727)	(0.0001587)	(0.0001769)
num varieties^2	-0.0001859	-0.0001274	-0.0001026	-0.0001028
	(1.129e-05)	(7.16e-06)	(4.656e-06)	(4.247e-06)
avg. size^2	-1e-05	-6.301e-06	-4.762e-06	-5.255e-06
	(7.395e-07)	(5.393e-07)	(4.217e-07)	(3.81e-07)
nest parameter	0.99	0.99	0.99	0.99
	(0.004894)	(0.00514)	(0.004873)	(0.006927)

Notes: Table reports coefficients from demand estimation where a market is a county-year-income group. We only use sales from dollar stores, discount stores, club stores, drug stores, convenience stores, and grocery stores. Data is based on consumer panel microdata for years 2008-2019. Income rank 1 is the lowest average income, income rank 4 has the highest average income.

Furthermore, the model allows us to quantify how households trade-off across product attributes. Specifically, under this model, an average household would be willing to give up 7% of varieties in order to purchase a bundle with 1% lower prices.

Dollar stores carry a limited selection of product groups; moreover, by ranking the product group by their fixed effect coefficients, the demand estimates provide intuition as to which product groups are more or less preferred by customers. Thus, if the dollar store were to select product groups entirely based on demand, then — under this model — the dollar store would likely select the product groups that provide most utility to consumer. We compare product group fixed effects from the demand estimates, detailed in Table A.10, to the groups purchased at the dollar store, detailed in Figure 2. We find little overlap between these two groups, indicating significant supply-side considerations for how dollar stores choose their product selection.

It has remained an open question as to how dollar stores have proliferated so successfully across the United States. We consider price, variety, size, and product selection as possible demand-side drivers of dollar store proliferation. We also consider convenience in Appendix A.3. Perhaps unsurprising given its name, we find that the demand-side driver of dollar store expansion is the dollar stores' low price point.

Finally, the demand estimation provides a different method to consider whether a consumer would be likely to switch from certain products at the grocery store to other products at the dollar store. That is, we compare the elasticity of substitution within nest to the elasticity of substitution across nests.[31] Since the nest parameter is .99, we find that the elasticity of substitution is much higher within nest than across nests; almost all of the substitution patterns occur within-nest or within store type. Thus, demand estimates suggest that price changes have a much larger effect on substitution patterns within store type than price changes across store types. In the context of policy questions about the dollar store, product groups like fresh produce at the grocery store are poor substitutes for product groups like snacks at the dollar store relative to other products at the grocery store.

In a similar vein, we also evaluate the elasticity of substitution across products at different store types directly.[32] Under our model, the elasticity of substitution between these goods depends on the price parameter and the share of each product. Since dollar store shares are small at currently observed levels, changes in prices at the dollar store cause little substitution from the grocery store to the dollar store. Similarly, the empirical analysis shows that the largest and most persistent substitution patterns occur for dry goods, the most common dollar store product group, and the smallest for fresh produce, a less common dollar store product group.

7.3 Quantifying the value of the dollar store

To quantify the benefit of the first dollar store to the average household, we estimate the compensating variation, the dollar value which equates the utility of the average household in a zip with a dollar store and the utility of the same household in the same zip but without a dollar store.[33] To use consistent measures of expenditure throughout the paper, we measure the log compensating variation, which we interpret in percentage terms,

$$\log CV = \frac{1}{\alpha} \ln \left(\frac{\left(\sum_{j \in s^{dollar}} e^{\delta_j/(1-\lambda)}\right)^{(} 1-\lambda}{\sum_{s \in s^{no\ dollar}} \left(\sum_{j \in s} e^{\delta_j/(1-\lambda)}\right)^{(} 1-\lambda)} + 1 \right) \tag{12}$$

where δ_j is the average utility $\delta_j = -\alpha \log p_j + \beta v_j + \beta_2 v_j^2 + \eta m_j + \eta_2 m_j^2 + \psi_{g(j)}$.

[31] Under the nested logit model, all products are substitutes. Products are by definition closer substitutes within nest than across nests, and the degree of substitution varies with the nest parameter. Product substitutability can vary from a logit model ($\lambda \to 0$), where all products have the same elasticity of substitution, to where the substitution within-nest is much higher than the substitution across nest, and all substitution occurs within the same nest ($\lambda \to 1$).

[32] Under the nested logit model, the elasticity of substitution between the share of good j and the price of good k is $\alpha p_k s_k$ for products across nests and $\alpha p_k s_k/(1-\lambda)$ for products within the same nest. Since our price parameter is expressed in logs, the elasticity of substitution under our model is αs_k.

[33] This is a standard measure of welfare for valuing new goods. For example, see Hausman and Leibtag (2007) in the context of Wal-mart's proliferation.

Qualitatively, household welfare increases after the first dollar store entry because the household choice set is increased. In practice, the welfare effects will be small, since the share of food products sold by the dollar store compared to other store types are small, as shown in Figure A.20. We compute the compensating variation using parameters from the demand estimation and the prices, characteristics, and number of trips from the Neilsen data. We use data in the year after the first dollar store entry to compute the compensating variation.[34]

We report the welfare estimates in Table 4.[35] We find that the benefit to the average household range from 11% for the average household in the top income rank to 17% for the average household in the bottom income rank. The average household is in rank 2 and spends 598$ per year on food expenditures, which translates into a 86$ value of the dollar store to the average household.

Table 4: Welfare Estimates

Variable	Income Rank 1	Rank 2	Rank 3	Rank 4
log CV	0.1678	0.1445	0.1245	0.115

Notes: Table reports coefficients from demand estimation where a market is a county-year-income group. We only use sales from dollar stores, discount stores, club stores, drug stores, convenience stores, and grocery stores. Data is based on consumer panel microdata for years 2008-2019. Income rank 1 is the lowest average income, income rank 4 has the highest average income. Standard errors are below estimates. Standard errors were estimated using 1000 bootstrap iterations.

8 Heterogeneity by Income and Retail Environment

Dollar stores disproportionately locate in low-income neighborhoods (as well as non-metro areas), as shown in Figure A.4, and, on the flip side, dollar store customers are disproportionately low income, as shown in Figure A.5. Since dollar stores strategically locate in specific areas, households in these locations might experience the dollar store differently than the average consumer.[36]

In this section, we investigate how the first dollar store entry affects heterogeneous populations. We repeat our main analysis, but allow for heterogeneity by income as well heterogeneity for

[34] Specifically, for each market we compute the compensating variation using demand coefficients and price, variety, and size data. To compute welfare, we average log compensating variation across markets. We compute welfare using the first period after dollar store entry. The procedure is repeated 1000 times.

[35] To obtain standard errors, we bootstrap the entire quantification, following Horowitz (2001). Specifically, we draw a random sample of households with replacement from the transactions data and re-estimate relative prices, characteristics, and demand characteristics. We then re-compute average log compensating variation.

[36] The media has often highlighted that dollar stores shoppers are disproportionately low income and locate disproportionately in rural towns.

different retail environments. We consider two definitions of sparse retail environments: non metro areas (to capture the effect of dollar stores in rural areas) and food deserts Additional dimensions of sparse retail environments include concentrated retail environments, or places where househoulds shop at relatively few retailers. For each dimension of heterogeneity, we focus on the potentially most vulnerable population: the lowest income group and the sparsest retail environments.

Then, we consider the effect of the first dollar store entry on expenditures, quantities, prices, and varieties. For expenditures, we consider both total, dollar store, and grocery store expenditures. We expect dollar store expenditures to increase more for low-income consumers, as well as as for consumers in sparse retail environments. To asses whether there maybe a competitive response from local retailers, we look at the effect on grocery expenditures and relative prices. Finally, we compute the effect on quantities and varieties to assess how heterogeneous effects compare to the average. For quantities, we consider the most policy-relevant department, fresh produce.

First, we consider heterogeneity by income. All income groups increase dollar store expenditures following the entry of the first dollar store, as shown in Figure A.21. However, households in the lowest income group react to the dollar store in a way which is different from other income groups and largely budget neutral. While other household groups react by reducing expenditures and varieties, the lowest income group barely reduces varieties and expenditures, and in a way that is not statistically significant, as shown in Figure A.25 and Figure A.21. Similarly, the lowest income group's grocery store expenditure, shown in Figure A.23, do not decline significantly following dollar store entry. Intuitively, households in the lowest income group are not using the dollar store to save money, but are using the dollar store to re-optimize their choices at (approximately) the same budget constraint.[37] Meanwhile, the lowest income relative price index (RPI) increases marginally and not significantly (by .01 percentage points), as shown in Figure A.24. Finally, the lowest income households do not change the ounces of fresh produce following the first dollar store entry, as shown in Figure A.26. Overall, this paints a picture of the average household in the lowest income group which continues to operate at its previous budget constraint, and re-optimizes utility but in a way that is budget neutral and in a way that does not change the quantity of fresh produce consumed.

Unlike for the lowest income households, the shopping patterns for households in food deserts and non-urban areas mirror the average trend (although results are less precise due to data availability). That is, households in sparse retail environments increase dollar store expenditures, with household in food deserts spending an additional 15$ per quarter at the dollar store, well above the average. Similar to the average, households in food deserts and non-urban areas reduce expenditures, grocery expenditures, and variety, and experience the same lack of price change and lack of change in the quantity of fresh produce.

[37] These seemingly disparate facts – that dollar store expenditures decrease while total expenditures do not change – is due to the fact that expenditures reduce marginally at other store types but in such a way that total expenditure is decreasing but not significantly.

9 Conclusion

With over 35,000 dollar stores in the United States, the dollar store format has been subject to scrutiny regarding its effects on consumers. Much of this scrutiny has centered around three related concerns regarding (1) the types of goods supplied and how these goods differ from those provided by traditional retail formats, (2) the effect of dollar stores on local retail competition and the consumer choice set, and (3) the effect of the dollar store on consumer welfare. This paper addresses these questions in the context of the first dollar store entry. We quantify the effect of the dollar store on households and local retailers, and investigates the mechanisms that drive the supply and demand response.

First, we document the types of goods dollar stores carry and how they differ from goods supplied by other store types. We show that, relative to other store types, dollar store goods are characterized by their low prices, small sizes, and few varieties and we show that dollar store entry introduces mostly dry goods into the market. Even per unit, the dollar store prices are the lowest amongst all other store types. This low-price result voids the concern that dollar stores exploit cash-strapped consumers by charging low prices per good but high prices per unit by only offering small sizes.

However, this concern illustrates the uniqueness of the dollar store format. While dollar stores offer simultaneously the cheapest prices and the smallest sizes, grocery stores tend to exhibit an inverse relationship between price and size. This inverse relationship allows grocery stores to price discriminate across consumers and increase effective prices by shrinking package sizes. Thus, one possible strength of the dollar store is the ability to reverse the traditional inverse relationship between size and price. It seems reasonable that dollar stores stock few varieties because they must guarantee a low price, and can only guarantee a low price per good and price per unit on a limited selection of goods.

In response to the first dollar store entry, we find a lack of supply-side response: while some store types (e.g. grocery stores) see declines in revenue following the first dollar store entry, stores do not change their prices and, on average, there is no grocery exit. On the demand side, in addition to not facing different prices at incumbent retailers, we show that households do not change the number of trips they take to other store types. Thus we conclude that the dollar store expands the household choice set.

Finally, we study the consumer response and quantify the welfare impact of the dollar store. We show that following the first dollar store entry, households reduce expenditures. This drop in expenditure is explained by a shift from a larger consumer bundle with higher prices but more variety to a smaller consumer bundle with lower prices but with less variety. The demand estimates suggest that this shift is driven by dollar store's low prices. We compute the value of the dollar store at 12% of food expenditure per year for the average household.

Our paper ties the literature on the expansion of non-traditional retail formats (such as Walmart, Sam's Club) to the literature on how households re-optimize in the face of large product assortments. We show the importance of re-optimizing over varieties and how non-traditional retail formats allow households to purchase new goods and goods more efficiently.

A Appendix

A.1 Tables

Table A.1: Product Size Effects of Dollar Store

Dependent Variable:	log(Size)
Variables	
Dollar Store	-0.2253
	(0.0027)
Discount	0.0487
	(0.0009)
Club	0.7244
	(0.0050)
Convenience	-0.2326
	(0.0060)
Drug	-0.1329
	(0.0053)
Fixed-effects	
county_product_month	Yes
Fit statistics	
Observations	483,480,789
R^2	0.82684
Within R^2	0.06925

Clustered (county_name) standard-errors in parentheses

Notes: Table reports coefficient from regressing log size of good on a store type variable. We only use sales from dollar stores, discount stores, club stores, and drug/convenience stores. Grocery stores are used as the reference group. Data is based on consumer panel microdata for years 2008-2018. We report coefficients with county by product module by month-year fixed effects. Standard errors are clustered at the county level.

Table A.2: Non-Food Retailers in the Retail Closing Instrument

Retail Chain	Bankruptcy Year	Retail Chain	Bankruptcy Year
Blockbuster	2010	Bon-Ton Stores	2018
Charlotte Russe	2018	Destination Maternity	2019
Fashion Bug	2013	Gymboree	2019
HH Gregg	2017	Hollywood Video	2010
KB Toys	2008	Loehmann's	2014
Mattress Firm	2018	Mervyn's	2008
Movie Gallery	2010	Radioshack	2017
Rue 21	2017	Shopko	2019
Sports Authority	2016	Tweeter	2007

Table A.3: Product Groups In Demand Estimation

Product Groups in Demand	
Bread and Baked Goods	Candy
Carbonated Beverages	Cereal
Cheese	Coffee
Cookies	Eggs
Fresh Produce	Ice Cream, Novelties
Juice, Drinks	Meat
Milk	Prepared Food
Snacks	Snacks – Other
Soft Drinks Non-Carbonated	Soup
Vegetables: Canned, Dried, and Frozen	Yogurt

We include the top 20 product groups by expenditure as our "inside goods" in our demand estimation. The remaining product groups comprise the outside good.

Table A.4: Product Groups Combined in Demand Estimations

These Product Groups	are Comprised of these Product Groups
Meat	Fresh Meat
	Packaged Meats - Deli
	Unprep Meat/Pultry/Seafood-frozen
Prepared Food	Prepared Food-Ready-To-Serve
	Prepared Food - Frozen
	Prepared Food - Dry Mixes
Juice, Drinks	Juice, Drinks - Canned, Bottled
	Juice, Drinks - Frozen
Vegetables - Canned, Dried, Frozen	Vegetables – Canned
	Vegetables and Grains - Dried
	Vegetables - Frozen
Snacks - Other	Snacks, Spreads, Dips - Dairy
	Pizza/Snacks/Horse Devours-Frozen
	Pudding, Desserts - Dairy

We combine similar product groups are combined into the same product group.

Table A.5: Conversion between counts and food weights in the Neilsen HMS data

Product Group Description	Weight (oz)
Egg	1.7
Fresh Apple	5.7
Fresh Cauliflower	32
Fresh Tomato	6
Fresh Potato	7.5
Fresh Mushroom	2
Fresh Onion	11.09
Fresh Kiwi	4
Fresh Grapefruit	8
Fresh Oranges	4.6
Fresh Lettuce	10.58
Fresh Garlic	1.41

We convert counts into weights (ounces) for several products.

Table A.6: First Stage for Hausman, Retail Closing Instruments

Dependent Variable:	store_price		
	First Stage: Hausman	First Stage: Retail Closing	First Stage: All
Model:	(1)	(2)	(3)
Variables			
hausman	0.8023***		0.8024***
	(0.0240)		(0.0240)
num varieties	-0.0027***	-0.0047***	-0.0027***
	(0.0006)	(0.0010)	(0.0006)
average size	0.0017**	0.0053**	0.0017**
	(0.0008)	(0.0023)	(0.0008)
num varieties^2	1.25×10^{-5}***	2.12×10^{-5}***	1.25×10^{-5}***
	(4.15×10^{-6})	(6.92×10^{-6})	(4.15×10^{-6})
average size^2	-1.31×10^{-6}	-6.32×10^{-6}	-1.31×10^{-6}
	(1.27×10^{-6})	(3.97×10^{-6})	(1.27×10^{-6})
banklag 1		0.0032***	-0.0035***
		(0.0008)	(0.0008)
Fixed-effects			
product_group_descr	Yes	Yes	Yes
Fit statistics			
Observations	2,911,532	2,911,532	2,911,532
R^2	0.10609	0.01767	0.10610
F-test (1st stage)	288,011.1	13.289	144,015.2

Clustered (product_group_descr) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

First stage regressions of log store price (store price) on the other retailer prices instrument (hausman) and the retail closing instrument (banklag 1).

Table A.7: First stage of retail closing instrument on demand entry.

Dependent Variable:	First dollar store entry to the zip code						
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Z_0	0.0329						0.0180
	(0.0202)						(0.0190)
Z_1		0.1306***					0.1700***
		(0.0092)					(0.0109)
Z_2			0.2153***				0.2514***
			(0.0082)				(0.0112)
Z_3				0.2741***			0.3077***
				(0.0113)			(0.0129)
Z_4					0.0309*		-0.0121
					(0.0185)		(0.0225)
Z_5						0.1460***	0.0508***
						(0.0147)	(0.0133)
Fixed-effects							
as.factor(panel_year)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics							
Observations	42,692	42,692	42,692	42,692	42,692	42,692	42,692
R^2	0.34928	0.35628	0.36974	0.37492	0.34919	0.35228	0.41272

Clustered (zip5) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.8: First stage of retail closing instrument on demand entry.

Dependent Variable:	Dollar Entry				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Z_1	0.0069***	0.0069***	0.0067***	0.0064***	0.0048***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0017)
Z_2	0.0132***	0.0132***	0.0130***	0.0124***	
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	
Z_3	0.0056***	0.0056***	0.0049***		
	(0.0015)	(0.0015)	(0.0015)		
Z_4	0.0120***	0.0118***			
	(0.0026)	(0.0025)			
Z_5	0.0016				
	(0.0025)				
Fixed-effects					
as.factor(zip5)	Yes	Yes	Yes	Yes	Yes
as.factor(time)	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	507,248	507,248	507,248	507,248	507,248
R^2	0.03716	0.03716	0.03710	0.03708	0.03694

Clustered (zip5) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.9: Placebo: regression of first stage of retail closing instrument on demand entry on stores of the wrong store size.

Dependent Variable:	opening				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Z_1	-0.0076***	-0.0072***	-0.0072***	-0.0070***	-0.0064***
	(0.0021)	(0.0021)	(0.0021)	(0.0021)	(0.0020)
Z_2	-0.0093***	-0.0091***	-0.0090***	-0.0089***	
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	
Z_3	-0.0049***	-0.0047***	-0.0046***		
	(0.0018)	(0.0018)	(0.0018)		
Z_4	-0.0019	-0.0017			
	(0.0018)	(0.0018)			
Z_5	-0.0038**				
	(0.0018)				
Fixed-effects					
as.factor(zip5)	Yes	Yes	Yes	Yes	Yes
as.factor(time)	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	511,504	511,504	511,504	511,504	511,504
R^2	0.03674	0.03673	0.03673	0.03671	0.03667

Clustered (zip5) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.10: Demand Estimates: Product Group Fixed Effects

Product Group	Income Rank 1	Rank 2	Rank 3	Rank 4
BREAD AND BAKED GOODS	-0.8675	-0.8219	-0.7862	-0.7782
	(0.02279)	(0.02379)	(0.02268)	(0.03219)
CANDY	-1.582	-1.596	-1.596	-1.649
	(0.02415)	(0.02499)	(0.02344)	(0.03242)
CARBONATED BEVERAGES	-0.9808	-0.9502	-0.9274	-0.9733
	(0.02244)	(0.0235)	(0.02236)	(0.03074)
CEREAL	-0.6689	-0.6375	-0.6315	-0.6668
	(0.02459)	(0.0255)	(0.02401)	(0.03343)
CHEESE	-0.4676	-0.4661	-0.4567	-0.4843
	(0.02599)	(0.0269)	(0.02536)	(0.03602)
COFFEE	-0.691	-0.6522	-0.5997	-0.6179
	(0.02695)	(0.0281)	(0.02682)	(0.03725)
COOKIES	-0.9222	-0.8831	-0.8479	-0.8556
	(0.02564)	(0.02703)	(0.02598)	(0.0364)
EGGS	-0.2677	-0.2328	-0.2265	-0.2428
	(0.02431)	(0.0255)	(0.02408)	(0.03367)
FRESH PRODUCE	-0.5363	-0.4991	-0.5006	-0.5374
	(0.02434)	(0.02501)	(0.02342)	(0.03307)
ICE CREAM, NOVELTIES	-0.3585	-0.313	-0.2535	-0.2816
	(0.02599)	(0.02707)	(0.02598)	(0.03553)
JUICES, DRINKS	-0.8417	-0.7653	-0.7508	-0.7545
	(0.02417)	(0.02537)	(0.0238)	(0.033)
MEAT	-0.5973	-0.5552	-0.5211	-0.5194
	(0.0243)	(0.02523)	(0.02392)	(0.03394)
MILK	-0.6419	-0.5362	-0.5263	-0.5411
	(0.02573)	(0.02724)	(0.02582)	(0.03557)
PREPARED FOOD	-0.9143	-0.8553	-0.7989	-0.7835
	(0.02311)	(0.02427)	(0.02317)	(0.0329)
SNACKS	-1.167	-1.164	-1.139	-1.157
	(0.02417)	(0.02494)	(0.02357)	(0.03294)
SNACKS - OTHER	-0.338	-0.3366	-0.3259	-0.3585
	(0.02658)	(0.0276)	(0.02618)	(0.03635)
SOFT DRINKS-NON-CARBONATED	-0.8695	-0.7987	-0.7921	-0.8051
	(0.02912)	(0.03017)	(0.02826)	(0.03832)
SOUP	-0.6416	-0.5985	-0.5587	-0.5815
	(0.0246)	(0.02577)	(0.0244)	(0.03418)
VEGETABLES- CANNED, DRIED, FROZEN	-0.7175	-0.6793	-0.63	-0.6434
	(0.02321)	(0.02414)	(0.02314)	(0.03291)
YOGURT	-0.1619	-0.2026	-0.2333	-0.2837
	(0.0272)	(0.02815)	(0.02637)	(0.03682)

Notes: Table reports coefficients from demand estimation where a market is a county-year-income group. We only use sales from dollar stores, discount stores, club stores, drug stores, convenience stores, and grocery stores. Data is based on consumer panel microdata for years 2008-2019. Income rank 1 is the lowest average income, income rank 4 has the highest average income.

Table A.11: Demand Estimates: Robustness, With Trips to Each Store Type

Variable	Income Rank 1	Rank 2	Rank 3	Rank 4
price	-0.7338	-0.7075	-0.6521	-0.5381
	(0.01426)	(0.01119)	(0.009932)	(0.008674)
num varieties	0.0408	0.03376	0.02992	0.03025
	(0.001387)	(0.0008539)	(0.0006102)	(0.0005873)
avg. size	0.005659	0.004308	0.003835	0.003627
	(0.0001829)	(0.0001599)	(0.000149)	(0.0001619)
num varieties^2	-0.0002385	-0.0001522	-0.0001206	-0.0001206
	(1.711e-05)	(8.874e-06)	(6.195e-06)	(5.097e-06)
avg. size^2	-9.625e-06	-6.365e-06	-4.995e-06	-5.575e-06
	(6.947e-07)	(5.034e-07)	(3.97e-07)	(3.626e-07)
nest parameter	0.99	0.99	0.99	0.99
	(0.004972)	(0.004732)	(0.004591)	(0.006353)
trips grocery	0.01243	0.009566	0.0089	0.00579
	(0.0003882)	(0.0003129)	(0.0002589)	(0.00023)
trips dollar	-0.1112	-0.1586	-0.1945	-0.2559
	(0.002431)	(0.00565)	(0.006144)	(0.0103)
trips convenience	-0.1262	-0.1481	-0.1788	-0.1588
	(0.01087)	(0.008352)	(0.01532)	(0.01524)
trips club	-0.06835	-0.04259	-0.03077	-0.01643
	(0.001375)	(0.0008253)	(0.0005882)	(0.0003734)
trips discount	-0.009605	-0.01128	-0.0148	-0.0244
	(0.0004414)	(0.0003806)	(0.0003297)	(0.0004313)
trips drug	-0.1645	-0.1446	-0.2141	-0.254
	(0.01089)	(0.005289)	(0.008243)	(0.009512)

Notes: Table reports coefficients from demand estimation where a market is a county-year-income group. We only use sales from dollar stores, discount stores, club stores, drug stores, convenience stores, and grocery stores. Data is based on consumer panel microdata for years 2008-2019. Income rank 1 is the lowest average income, income rank 4 has the highest average income.

Table A.12: Demand Estimates: Robustness, Grouping trips to stores as a single variable

Variable	Income Rank 1	Rank 2	Rank 3	Rank 4
price	-0.7969	-0.7087	-0.5845	-0.348
	(0.01373)	(0.01033)	(0.008897)	(0.007148)
num varieties	0.05343	0.04443	0.04094	0.03827
	(0.0008821)	(0.0006176)	(0.0005427)	(6e-04)
avg. size	0.006692	0.005188	0.004805	0.005028
	(0.0001903)	(0.0001655)	(0.0001494)	(0.0001628)
num varieties^2	-0.0002132	-0.0001448	-0.0001273	-0.0001299
	(1.181e-05)	(7.106e-06)	(5.665e-06)	(5.317e-06)
avg. size^2	-1.124e-05	-7.863e-06	-6.615e-06	-7.891e-06
	(7.481e-07)	(5.331e-07)	(4.065e-07)	(4.066e-07)
nest parameter	0.99	0.99	0.99	0.99
	(0.004839)	(0.004998)	(0.004638)	(0.006361)
non-grocery trips	-0.0367	-0.03426	-0.0369	-0.03458
	(0.0004192)	(0.0003309)	(0.000352)	(0.000414)

Notes: Table reports coefficients from demand estimation where a market is a county-year-income group. We only use sales from dollar stores, discount stores, club stores, drug stores, convenience stores, and grocery stores. Data is based on consumer panel microdata for years 2008-2019. Income rank 1 is the lowest average income, income rank 4 has the highest average income.

Table A.13: Welfare Estimates: Robustness, Grouping trips to stores as a single variable

Variable	Income Rank 1	Rank 2	Rank 3	Rank 4
log CV	0.1678	0.1445	0.1245	0.115

Notes: Table reports coefficients from demand estimation where a market is a county-year-income group. We only use sales from dollar stores, discount stores, club stores, drug stores, convenience stores, and grocery stores. Data is based on consumer panel microdata for years 2008-2019. Income rank 1 is the lowest average income, income rank 4 has the highest average income.

A.2 Figures

Figure A.1: Time Series of Dollar Stores

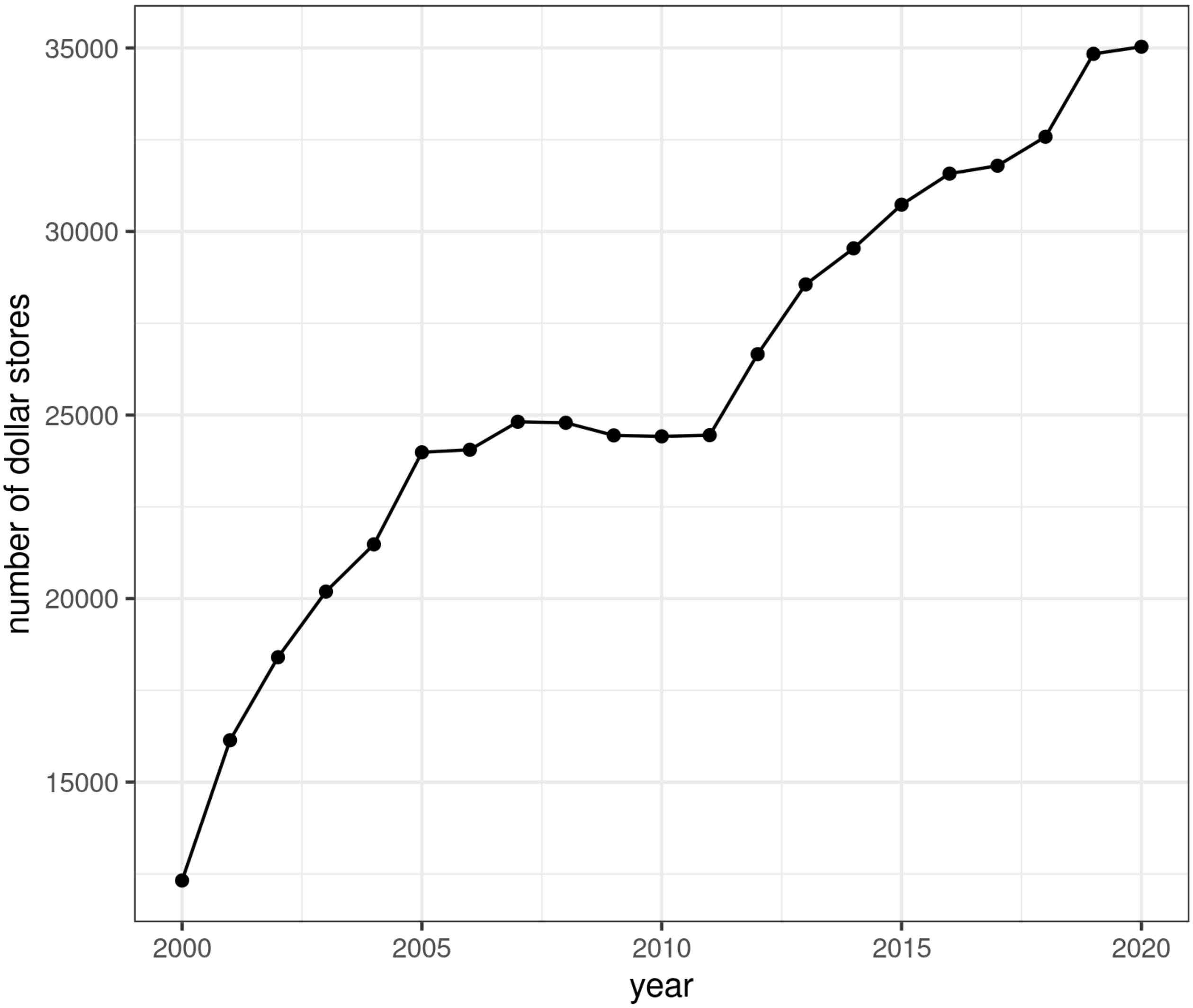
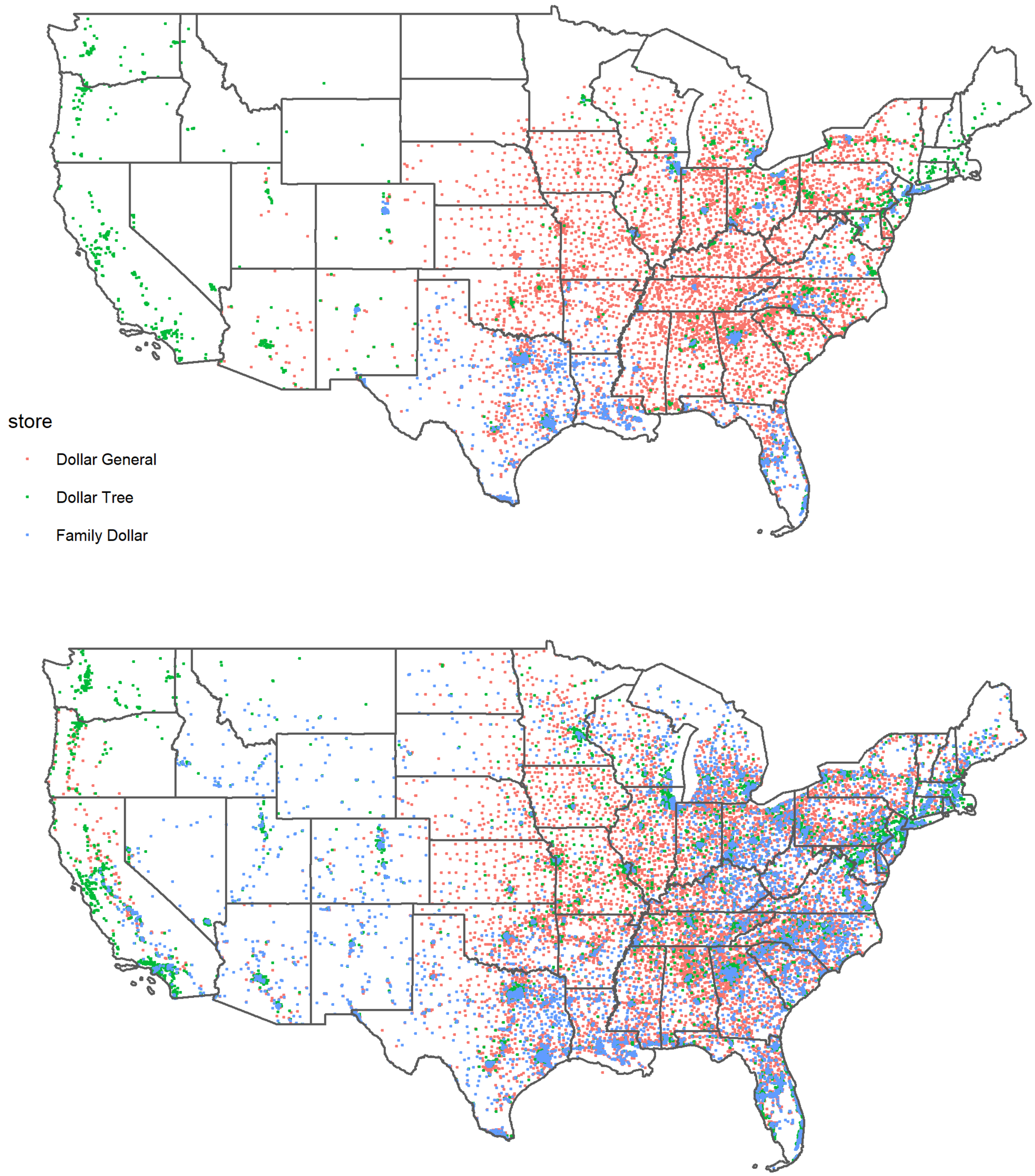
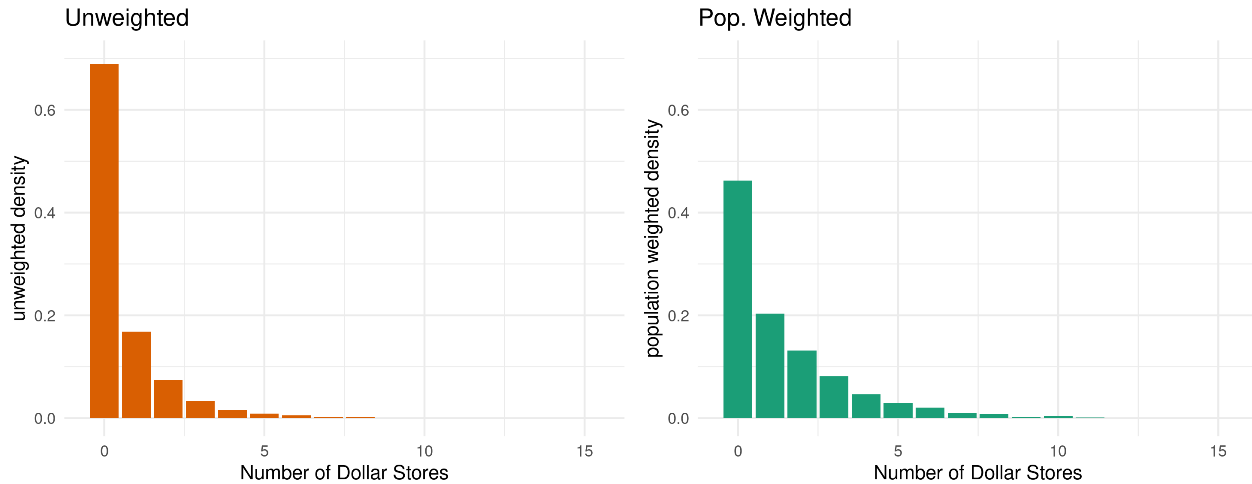


Figure A.2: Map of Largest Dollar Store Chains in 2008 and 2019



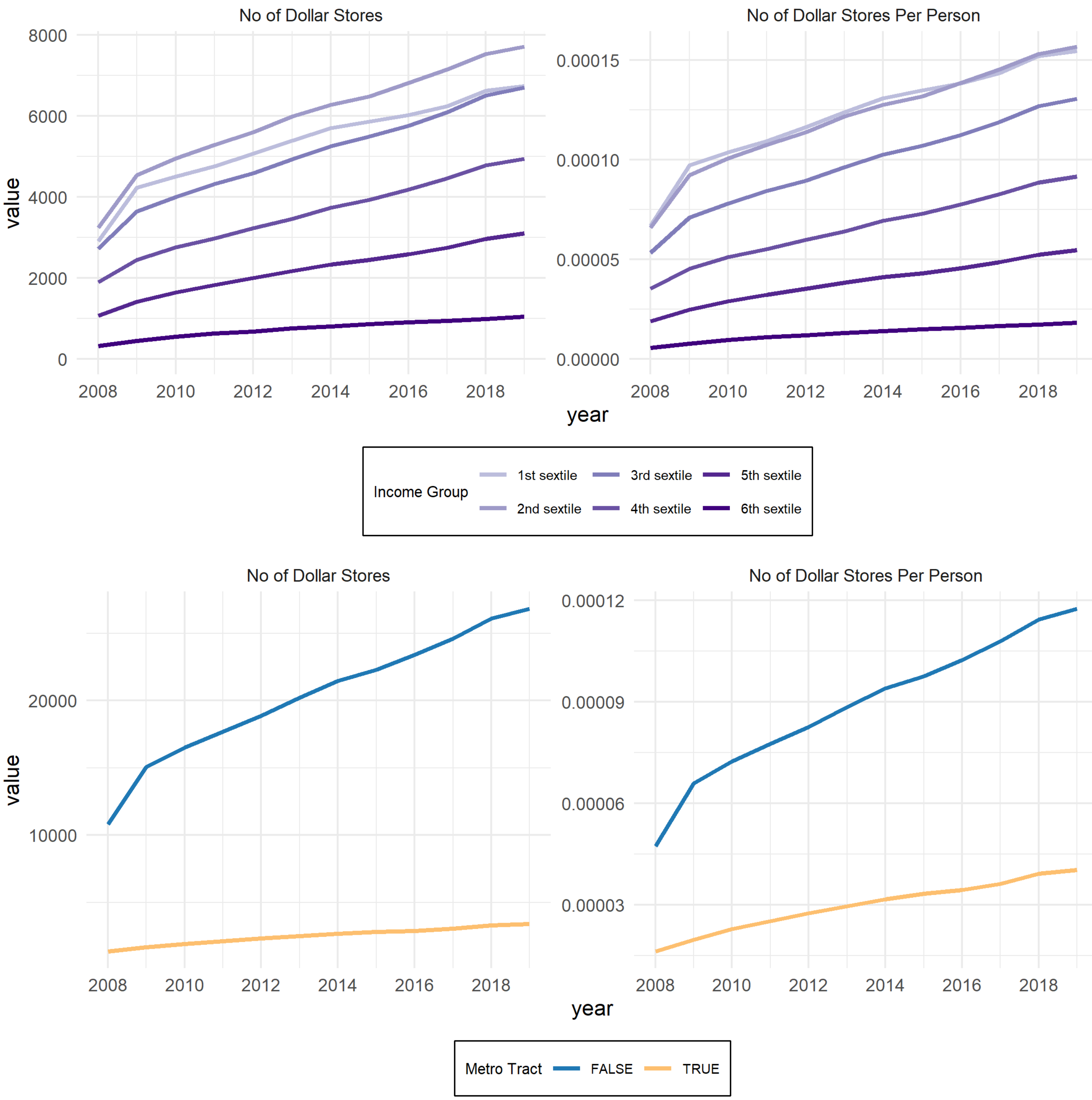
Notes: Dollar Tree, Dollar General, Family Dollar

Figure A.3: Dollar Store Density in Zip Codes



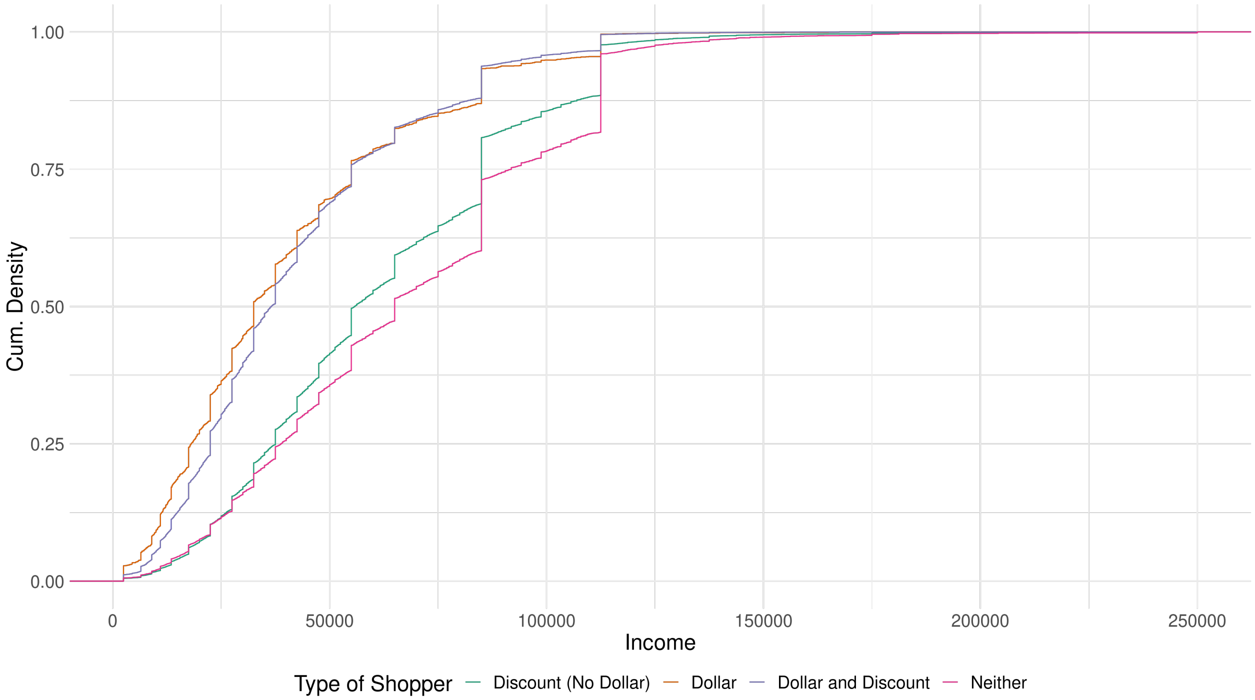
Notes: Plots current dollar store density unweighted (a) and weighted (b) by population. Each observation is a zip code. Most zip codes do not have a dollar store, and of the zip codes that do have dollar stores, most zip codes have one dollar store.

Figure A.4: Time Series of Dollar Store Growth by Income and Retail Environments



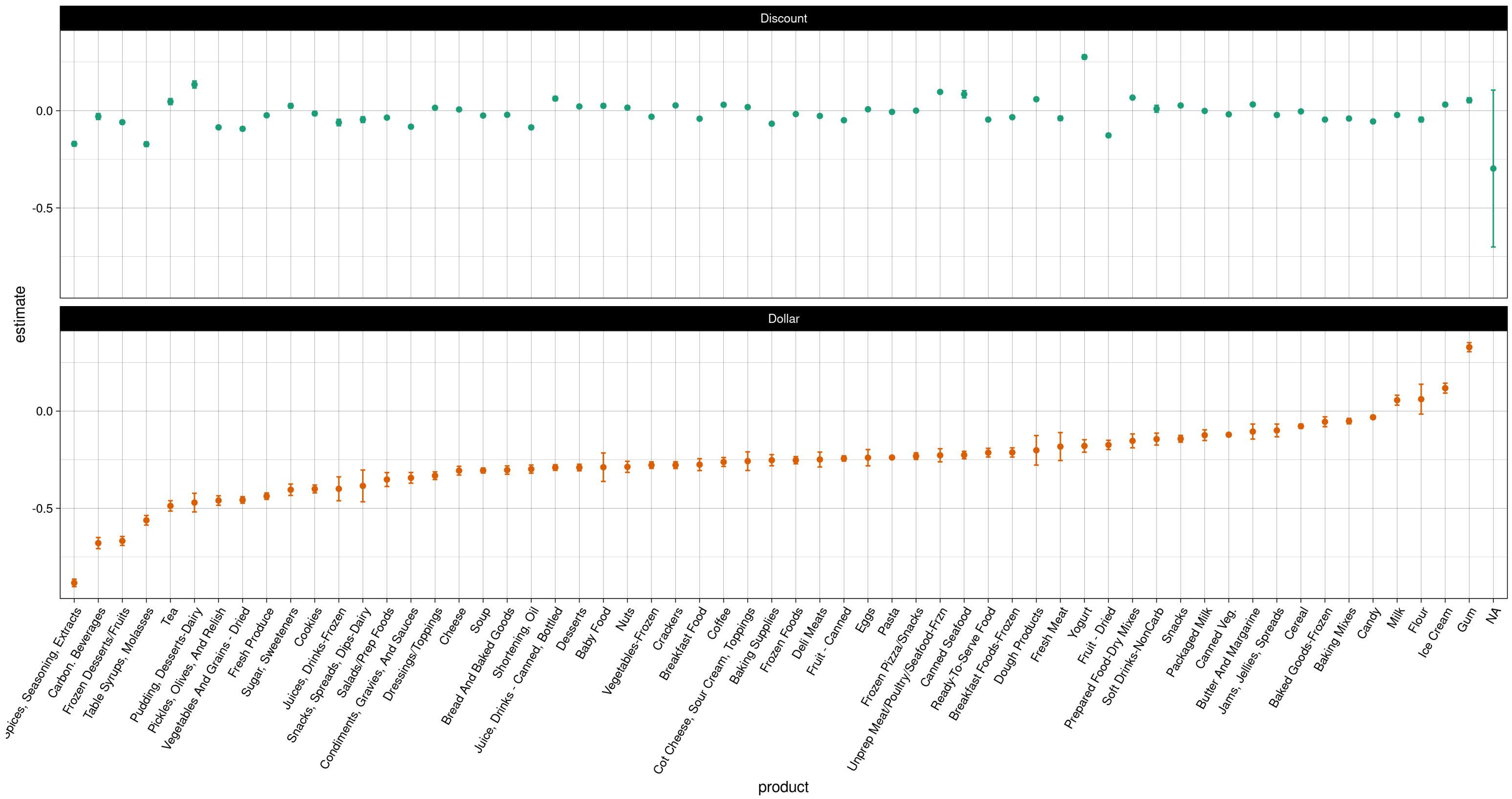
Notes: Growth of dollar stores over time by income and sparse retail environments

Figure A.5: Income Distribution by Where Households Shop



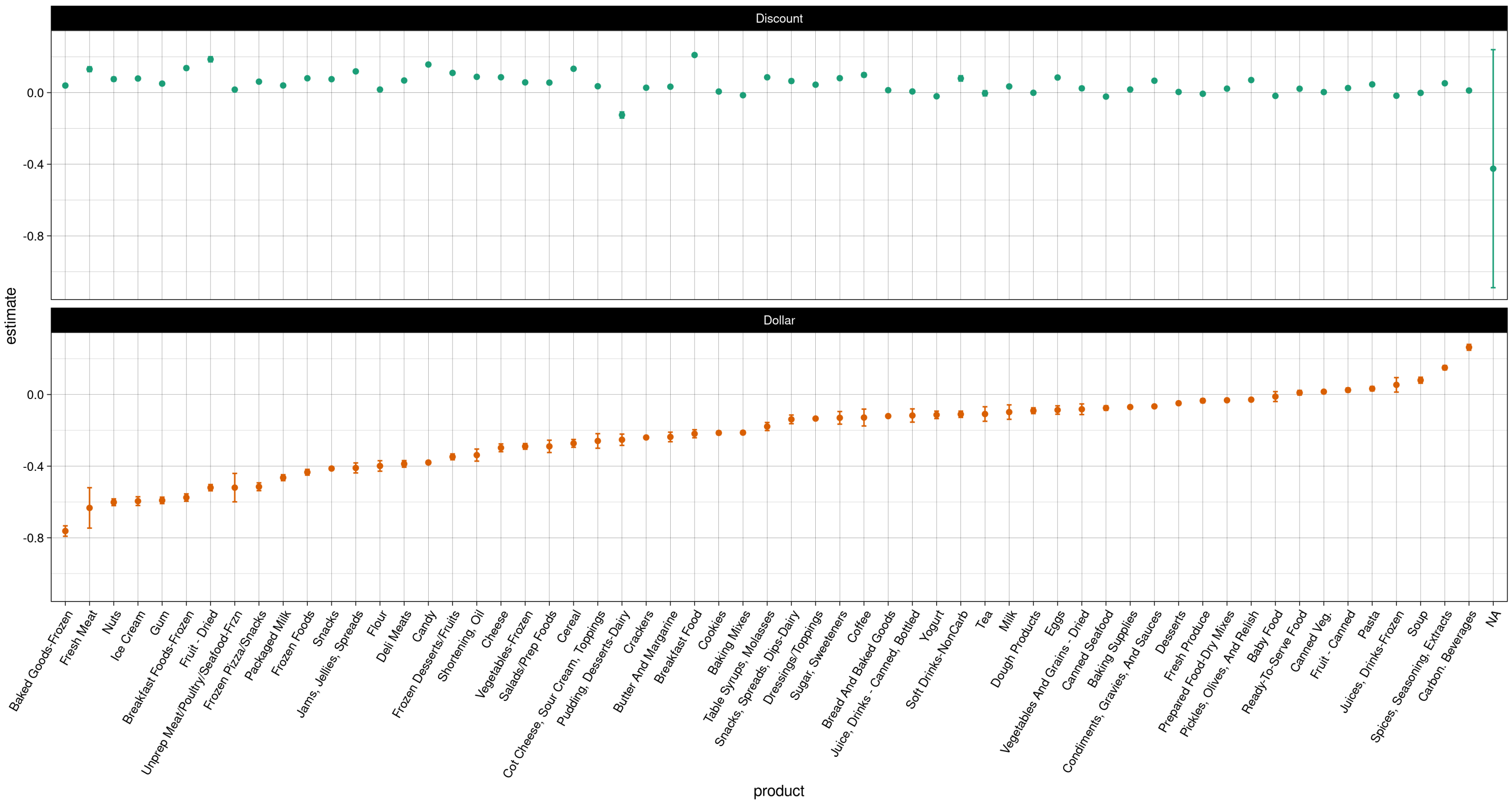
Notes: Figure reports CDF of the household income. The income distributions are broken out by shopper type. Dollar store shoppers are those that spent more than 5% of expenditures that year at dollar stores, and analogously for discount shoppers. Data is based on consumer panel microdata for years 2008-2019.

Figure A.6: Price Per Unit Effects By Product Group



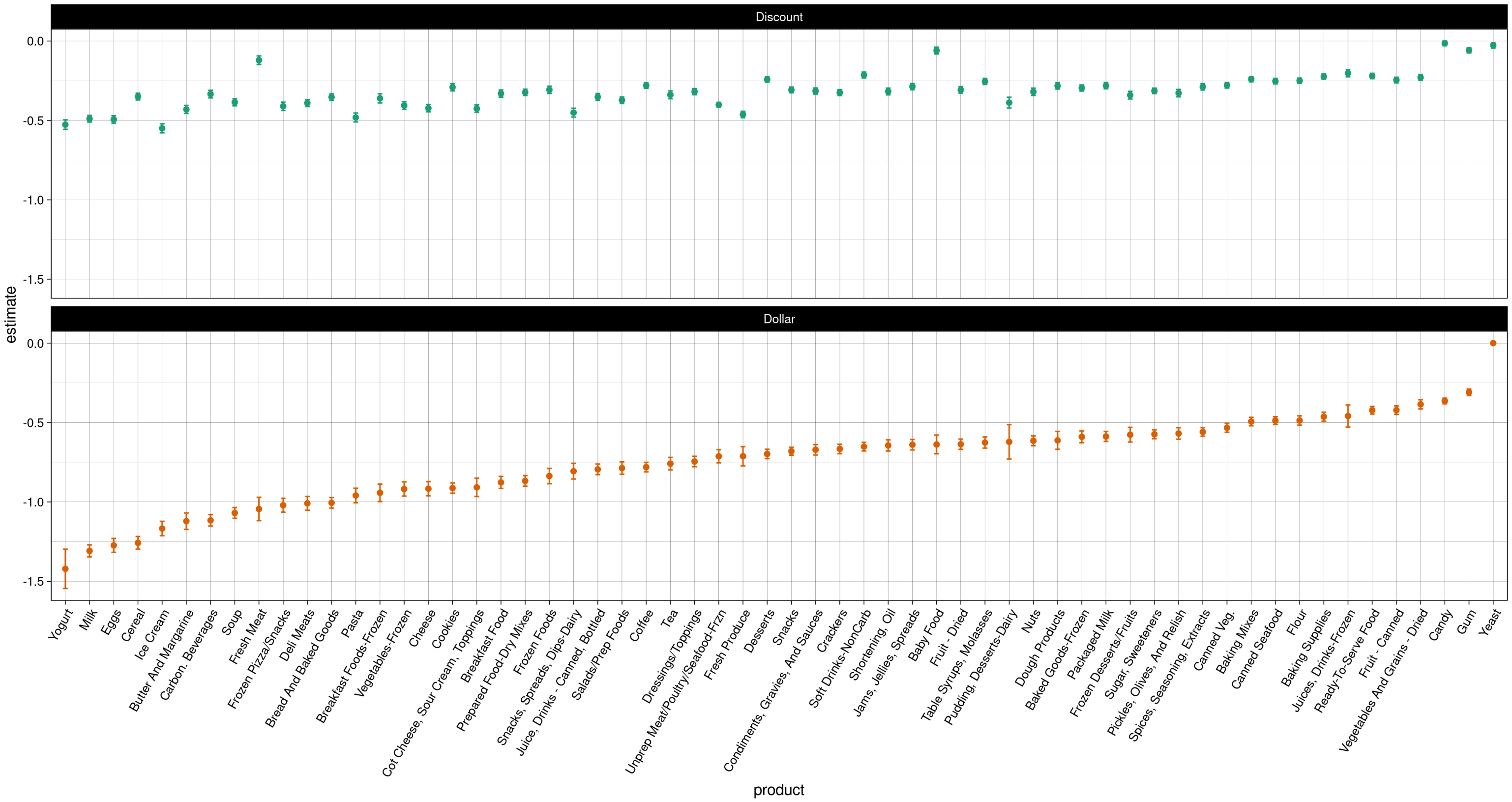
Notes: Figure reports coefficients and 95% confidence interval from regression of log price per unit of a good on store type, with county by product module by month-year fixed effects, broken out by product groups. We only use sales from dollar stores, discount stores, grocery stores, club/superstores, and drug/convenience stores. All stores other than dollar stores and discount stores are lumped into one group and used as the reference group for the regressions. Data is based on consumer panel microdata for years 2008-2018. Standard errors are clustered at the county level.

Figure A.7: Size Effects By Product Group



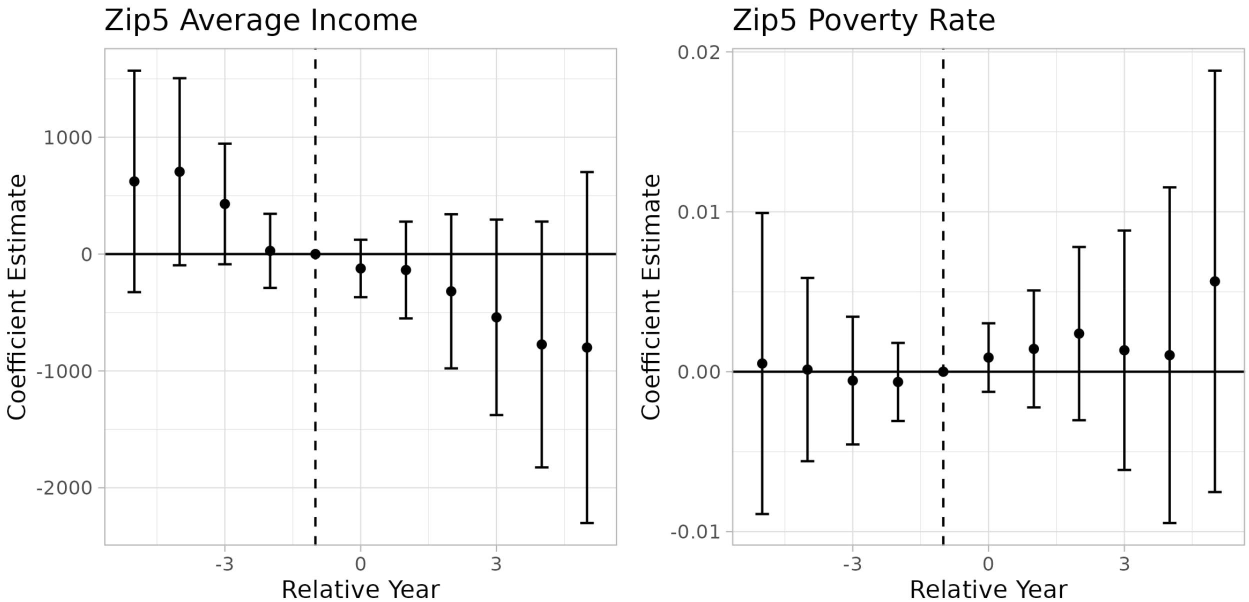
Notes: Figure reports coefficients and 95% confidence interval from regression of log size per unit of a good on store type, with county by product module by month-year fixed effects, broken out by product groups. We only use sales from dollar stores, discount stores, grocery stores, club/superstores, and drug/convenience stores. All stores other than dollar stores and discount stores are lumped into one group and used as the reference group for the regressions. Data is based on consumer panel microdata for years 2008-2018. Standard errors are clustered at the county level.

Figure A.8: Variety Effects By Product Group



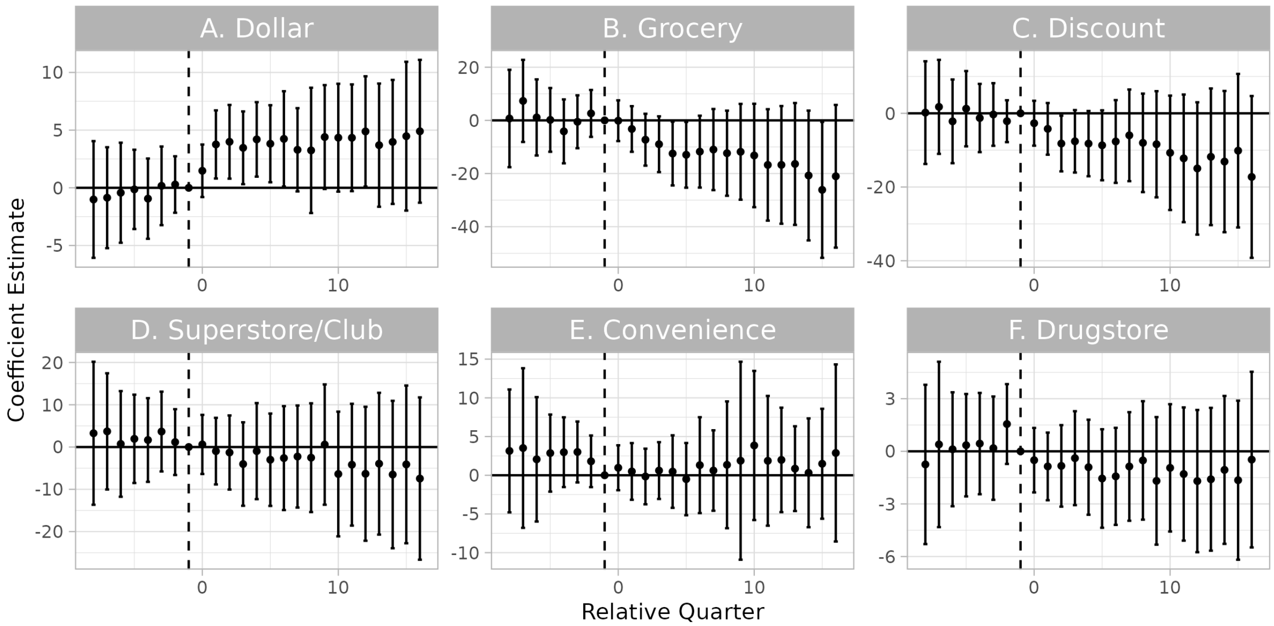
Notes: Figure reports coefficients and 95% confidence interval from regression of log size per unit of a good on store type, with county by product module by month-year fixed effects, broken out by product groups. We only use sales from dollar stores, discount stores, grocery stores, club/superstores, and drug/convenience stores. All stores other than dollar stores and discount stores are lumped into one group and used as the reference group for the regressions. Data is based on consumer panel microdata for years 2008-2018. Standard errors are clustered at the county level.

Figure A.9: Dollar Store Entry on Household Income



Notes: Figure reports event study estimates of impact on household income with 95% confidence intervals from Equation 5, using 5-year average zipcode income data using American Community Survey Data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.10: Dollar Store Entry on Expenditures by Store Type



Notes: Figure reports event study estimates of impact on store-level expenditures with 95% confidence intervals from Equation 5, using 2008-2018 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.11: Dollar Store Entry on Expenditure Spent in Each Department

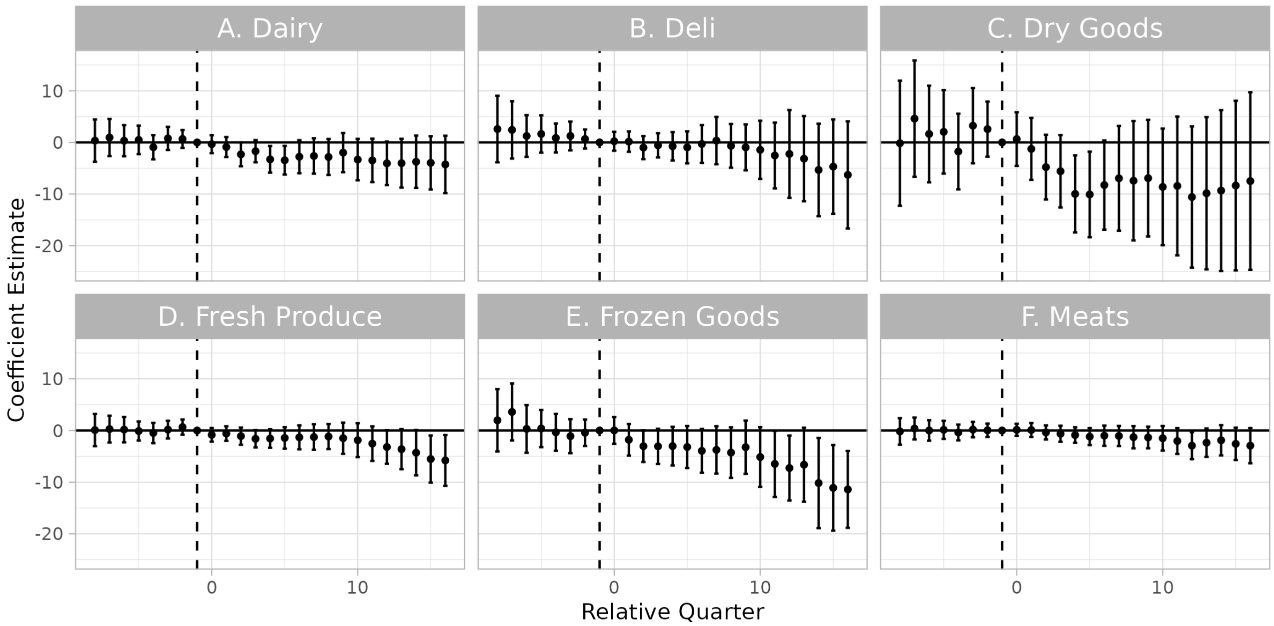
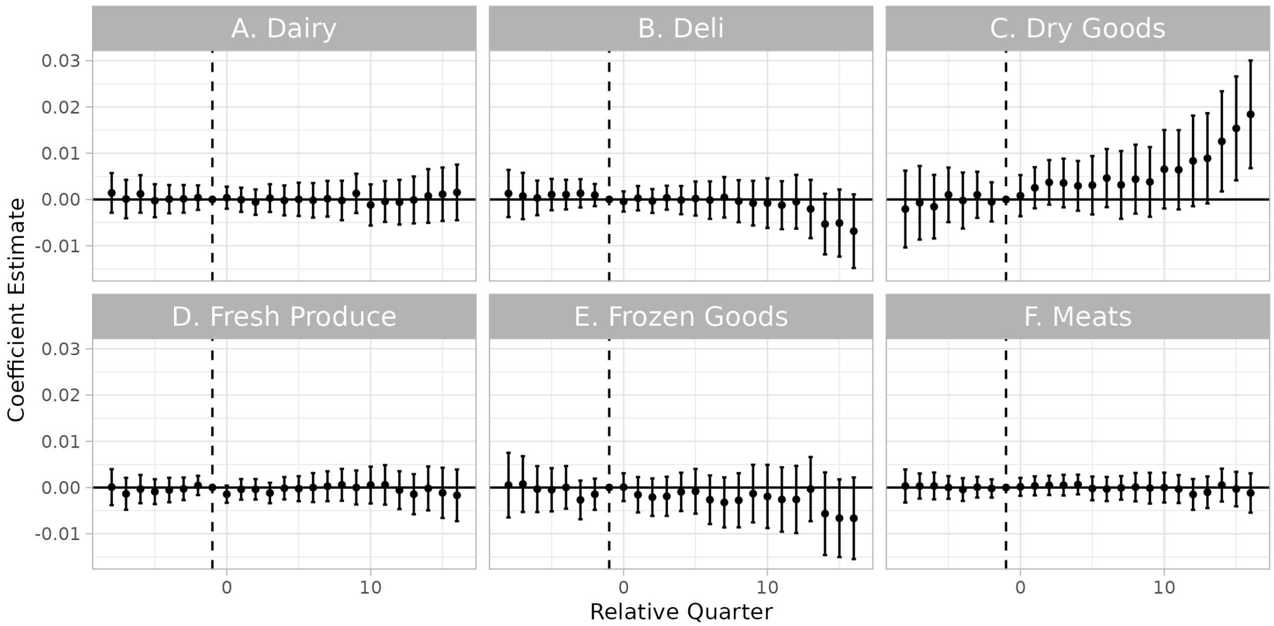


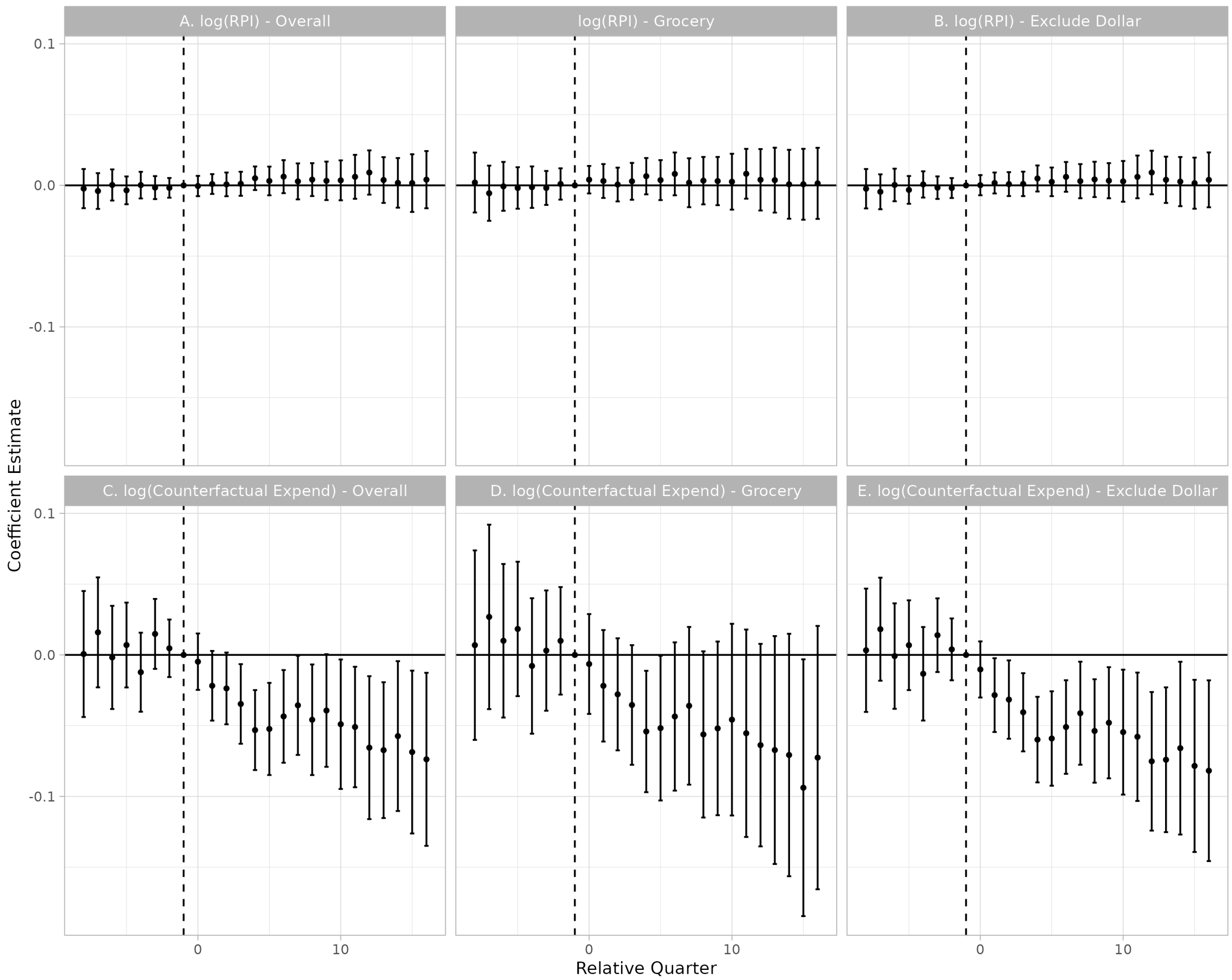
Figure A.12: Dollar Store Entry Fraction of Expenditure Spent in Each Department



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2018 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). The dependent variables here are different definitions for consumption bundle variety for the household at the quarter-year level. Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

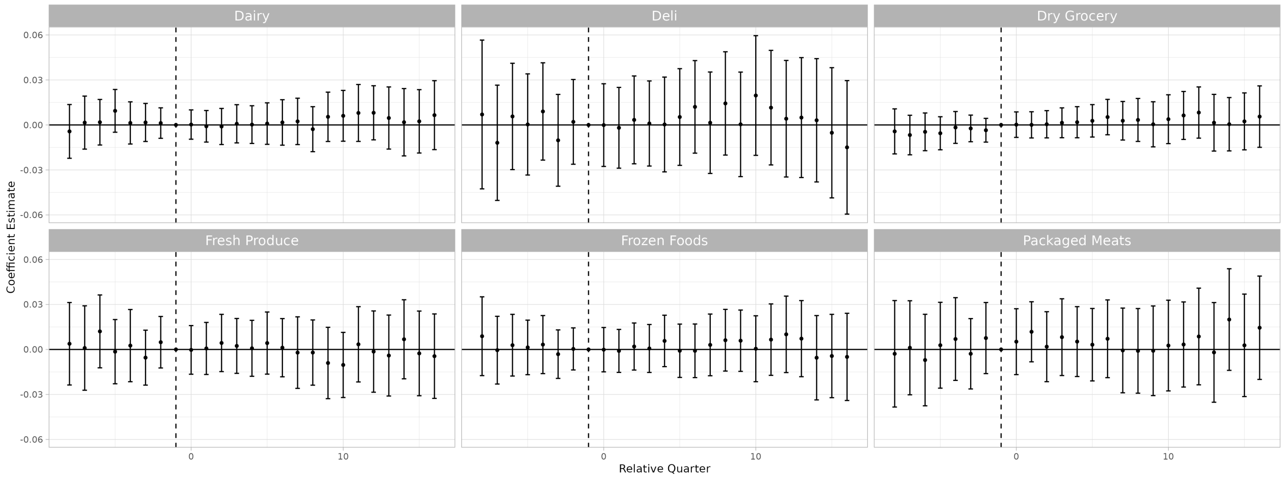
Figure A.13: Department Level Expenditure Shares by Store Type

Figure A.14: Effect of Dollar Store Entry on Log Relative Price Index and Counterfactual Expenditure



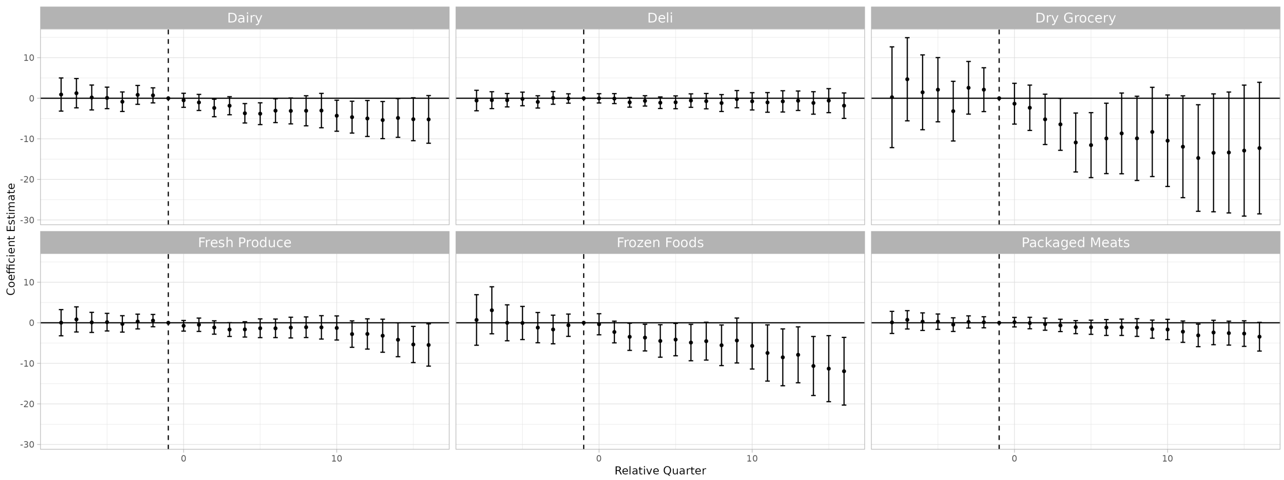
Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. The state is used to calculate the average reference price.

Figure A.15: Effect of Dollar Store Entry on Department-Level Relative Price Index



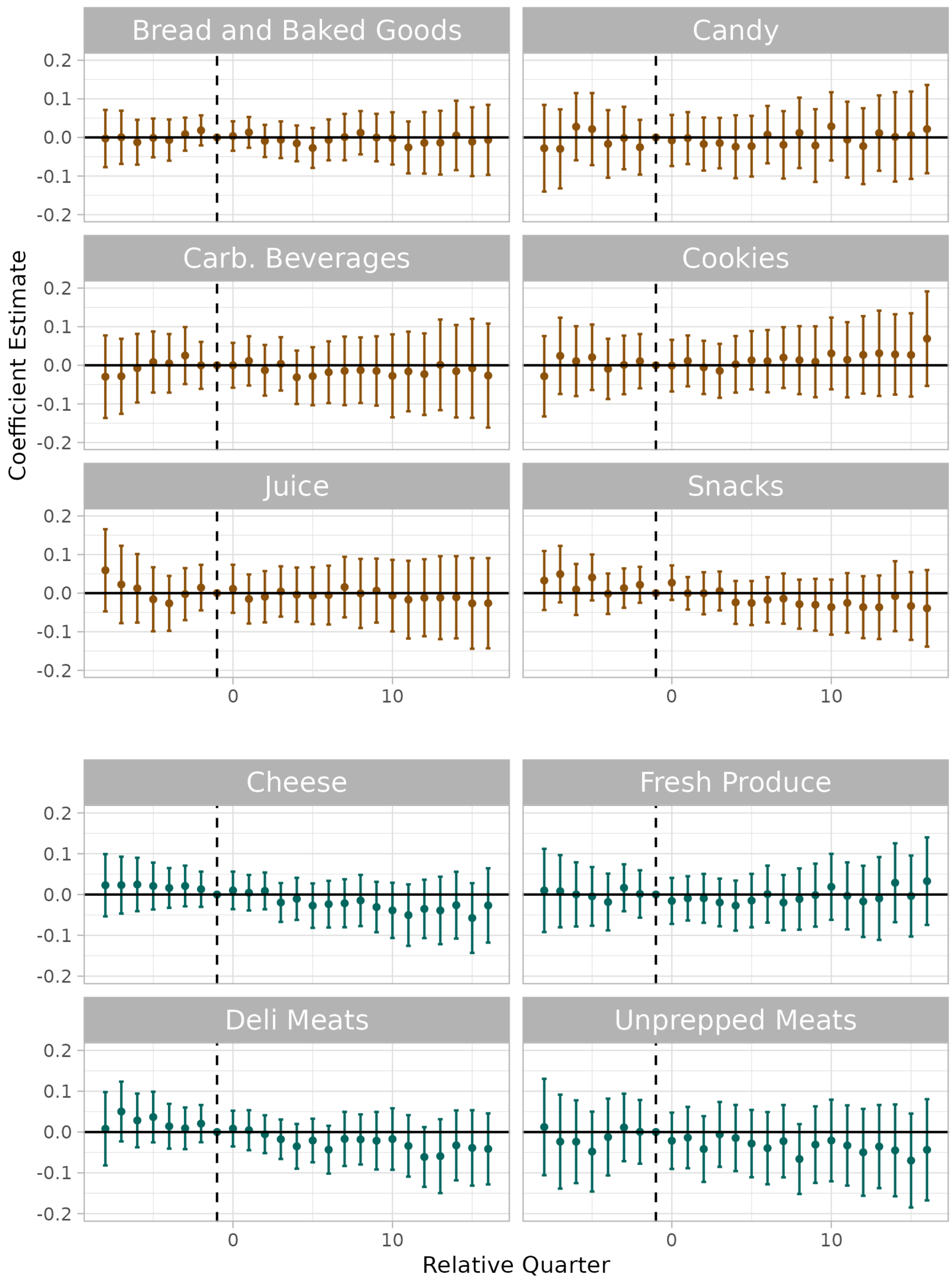
Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. The county is used to calculate the average reference price.

Figure A.16: Effect of Dollar Store Entry on Department-Level Counterfactual Expenditure



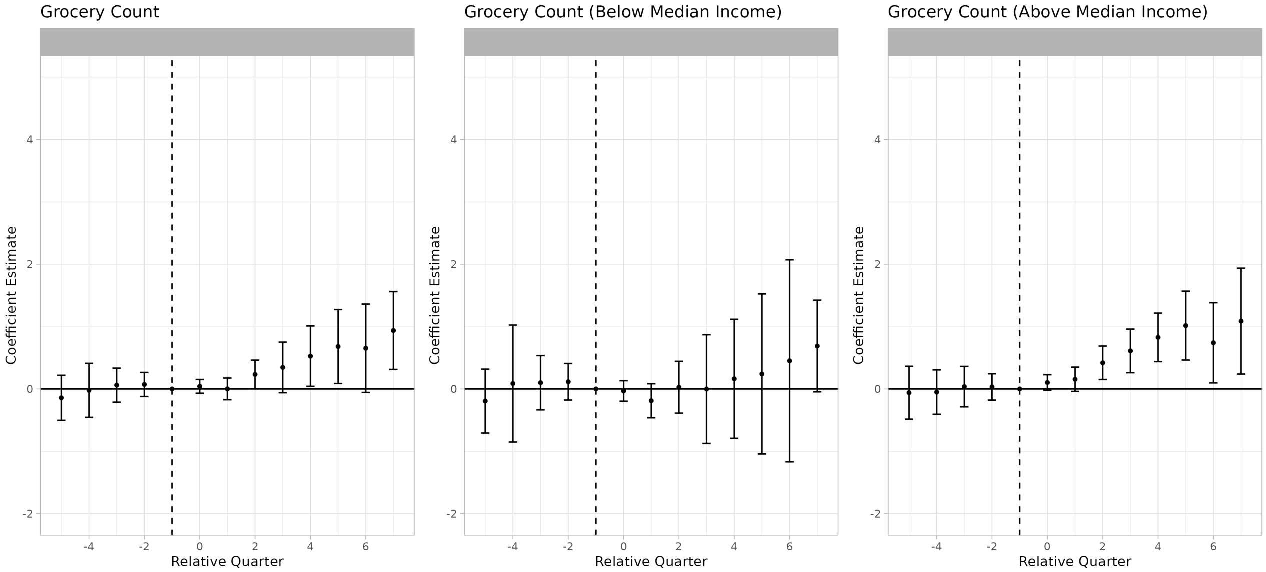
Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. The county is used to calculate the average reference price.

Figure A.17: Effect of Dollar Store Entry on Log Ounces of Each Product Group



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.18: Dollar Store Entry on Grocery Count: No Demographic Controls



Notes: Figure reports event study estimates with 95% confidence intervals from Equation 4, using 2008-2018 SNAP and ZBP data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). The figure reports the grocery count as the outcome variable. Errors are clustered at the zip code level.

Figure A.19: Retailer counts by year for each of the retailers that went bankrupt.

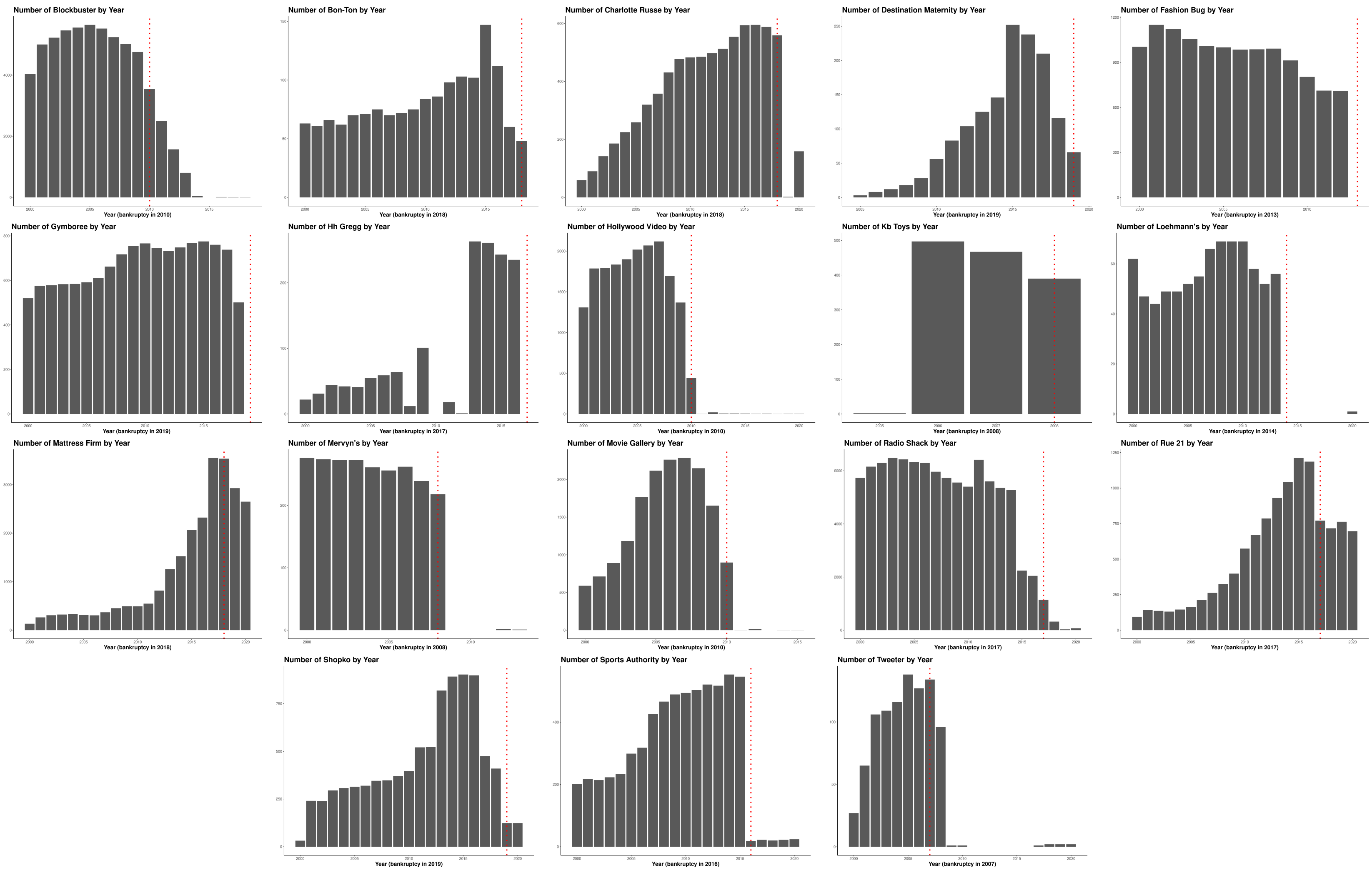
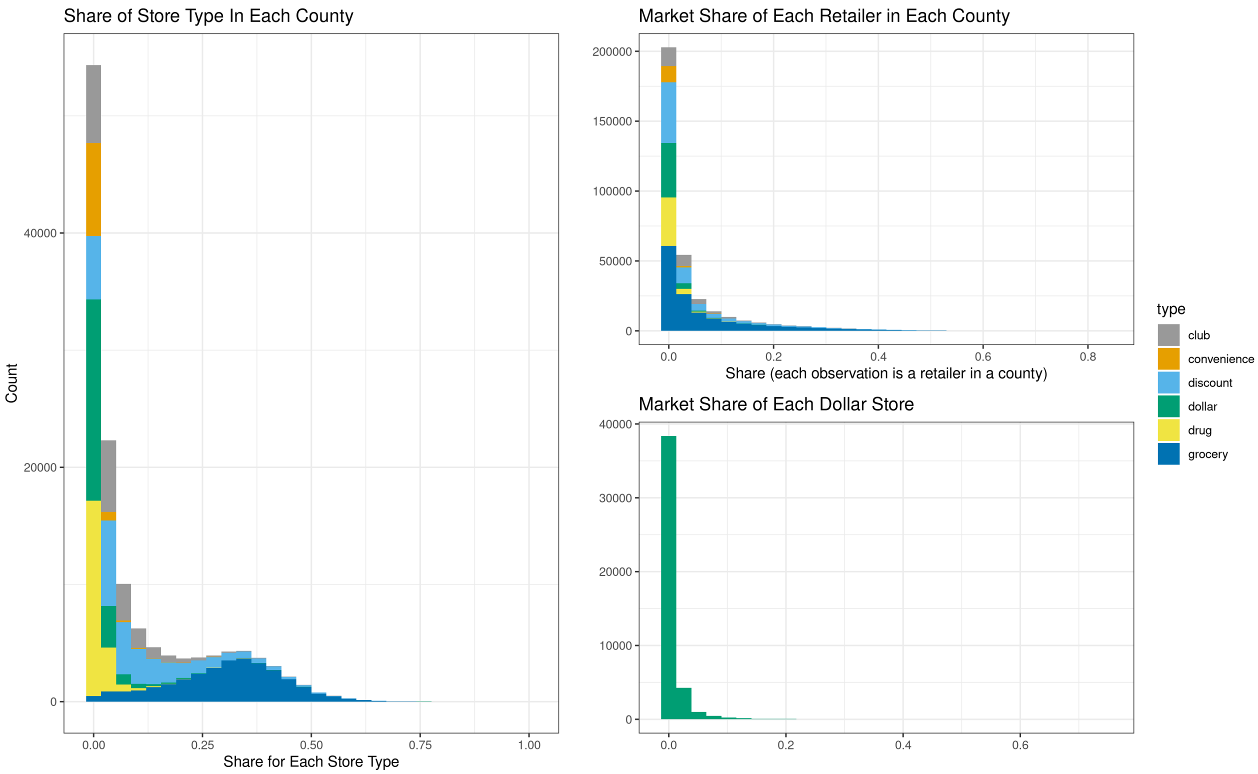


Figure A.20: Share of Each Store Type in each County-Year



Notes: Total quantity shares for each store type in each county, as measured as ounces of product from one store type as a fraction of the total ounces of food products sold in the county. Data comes from the Nielsen Homescan panel (2008-2019).

Figure A.21: Log Total Expenditure by Income

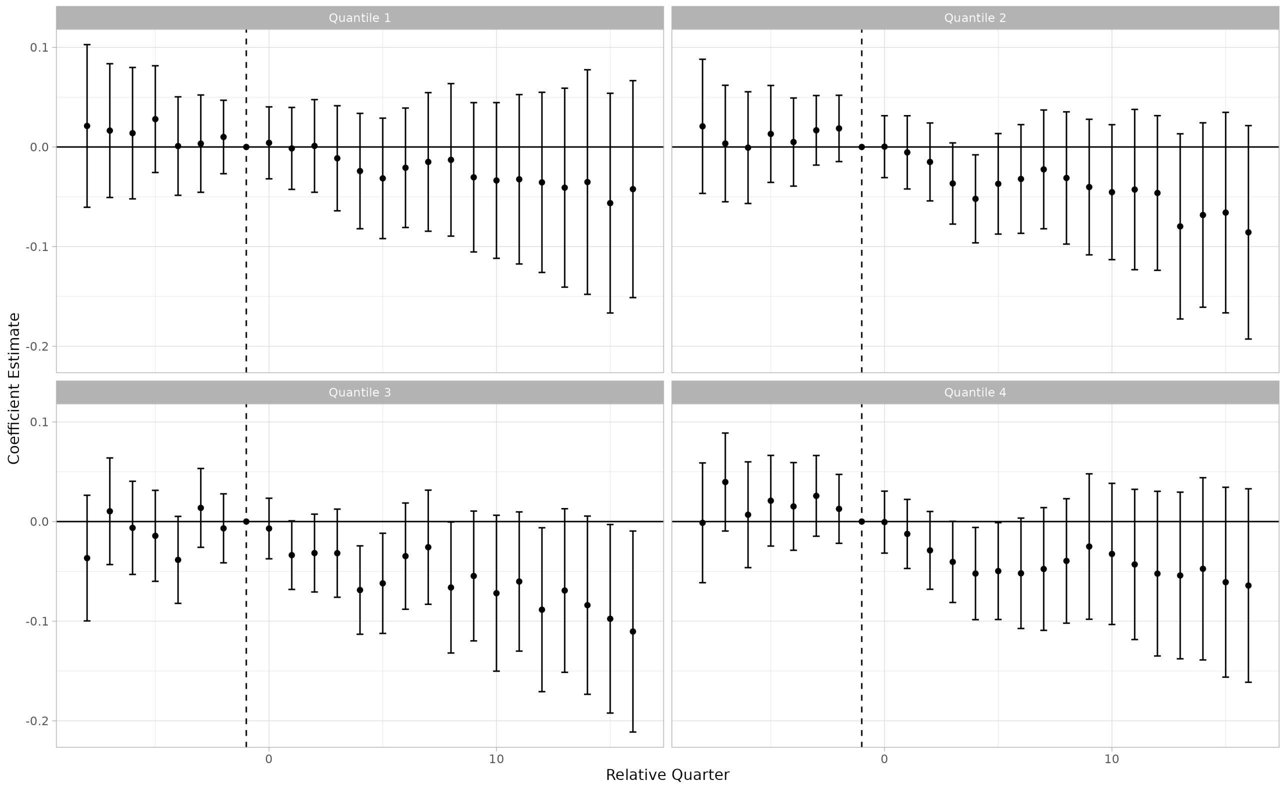


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. Income Rank 1 is the lowest income, Rank 4 is the highest income.

Figure A.22: Food Expenditure at the Dollar Store by Income

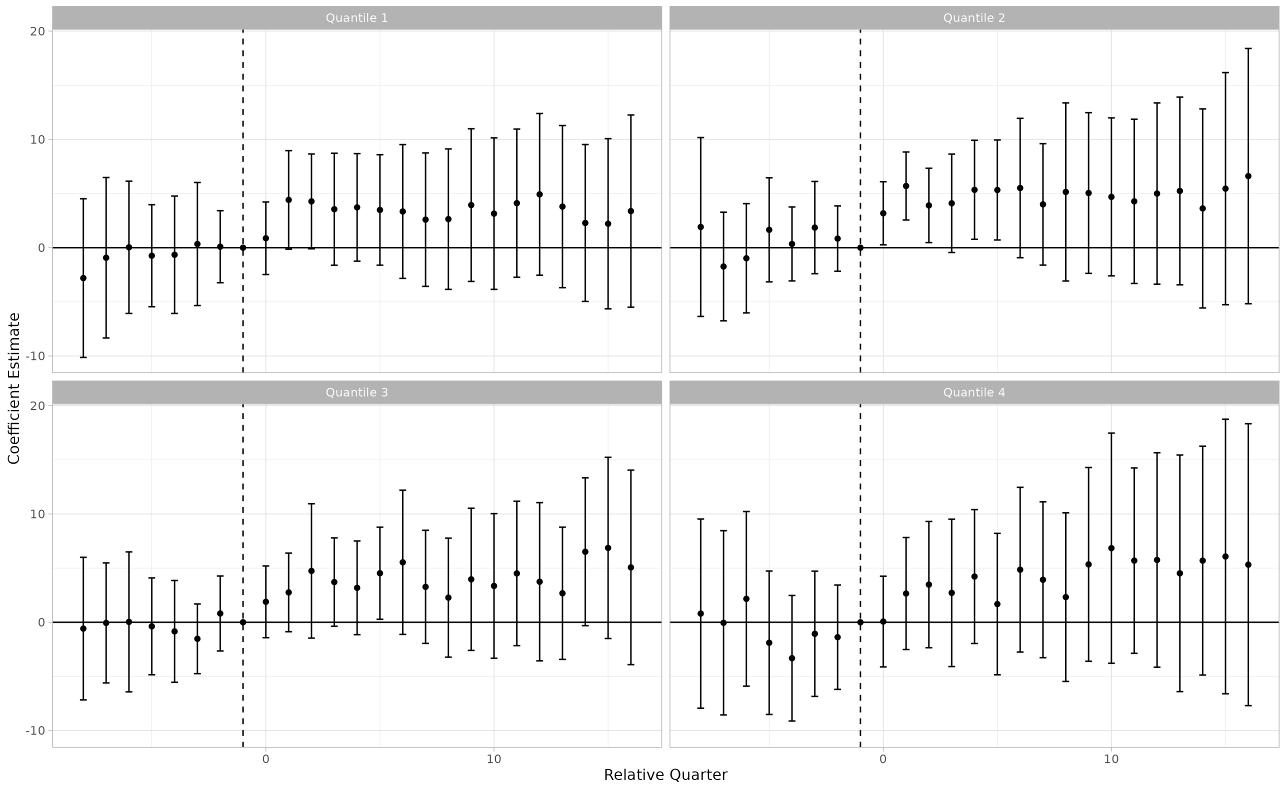


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. Income Rank 1 is the lowest income, Rank 4 is the highest income.

Figure A.23: Food Expenditure at the Grocery Store by Income

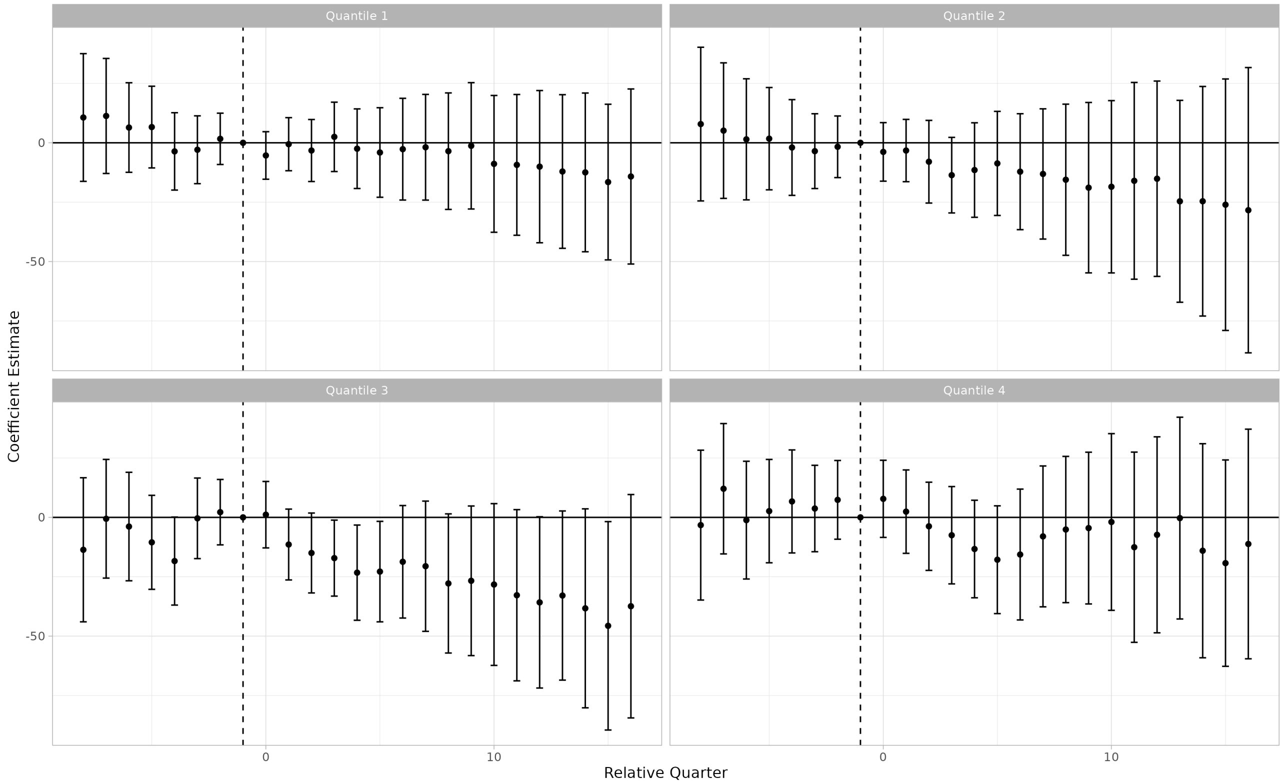


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. Income Rank 1 is the lowest income, Rank 4 is the highest income.

Figure A.24: Log Relative Price Index by Income

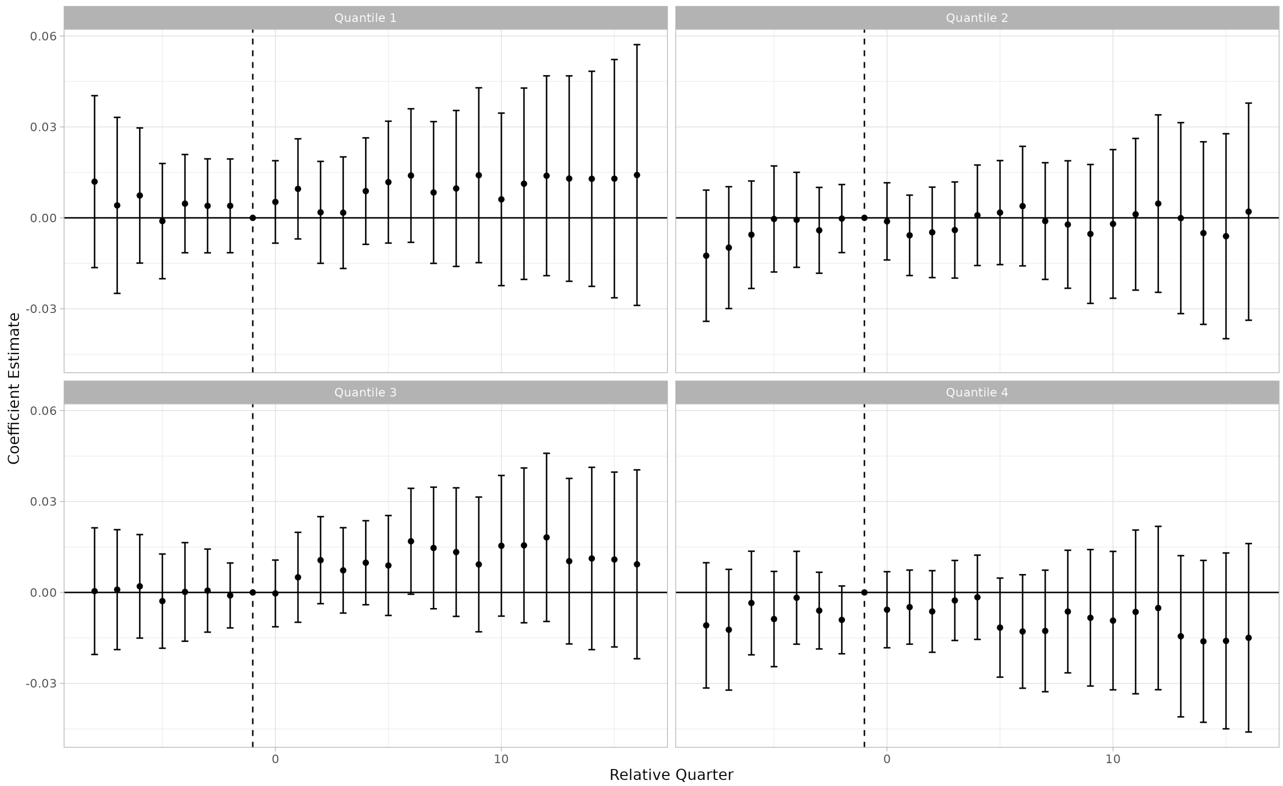


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. Income Rank 1 is the lowest income, Rank 4 is the highest income.

Figure A.25: Log Number of Unique Varieties by Income

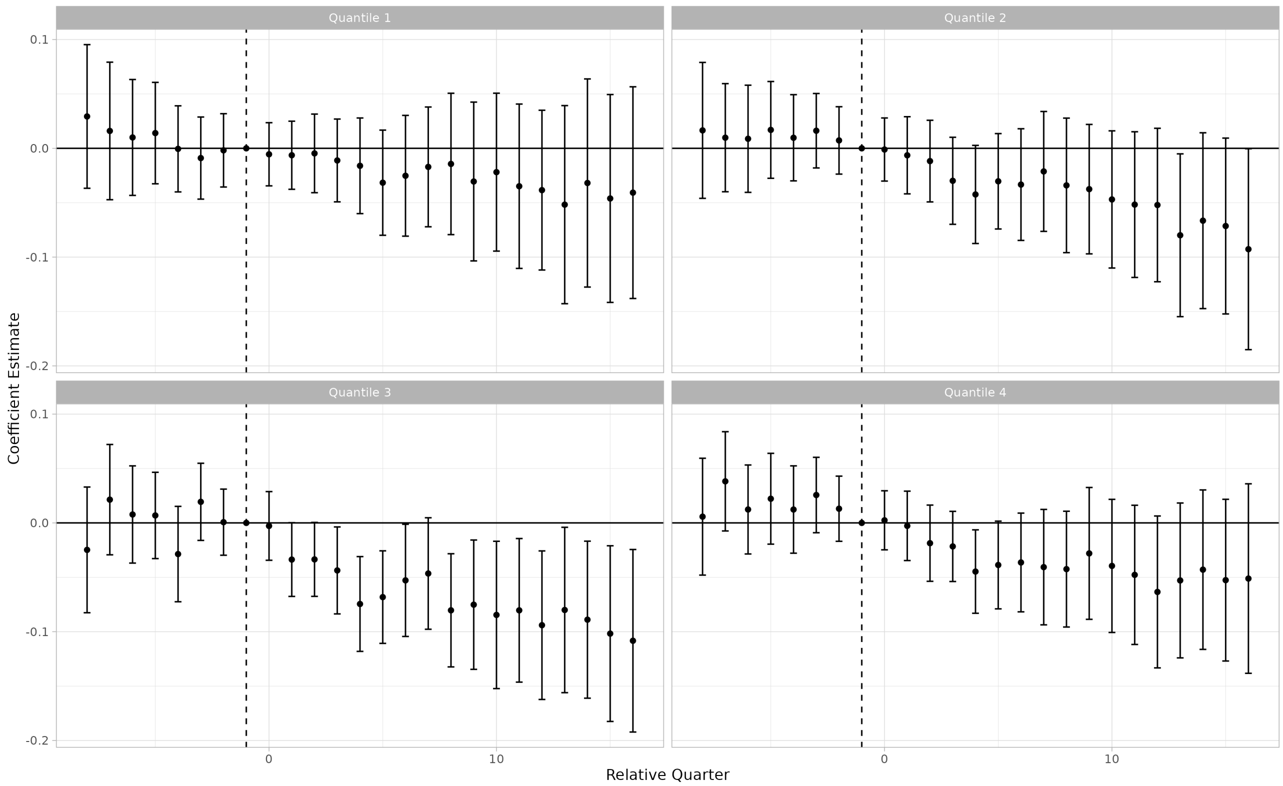


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. Income Rank 1 is the lowest income, Rank 4 is the highest income.

Figure A.26: Log Quantity of Fresh Produce by Income

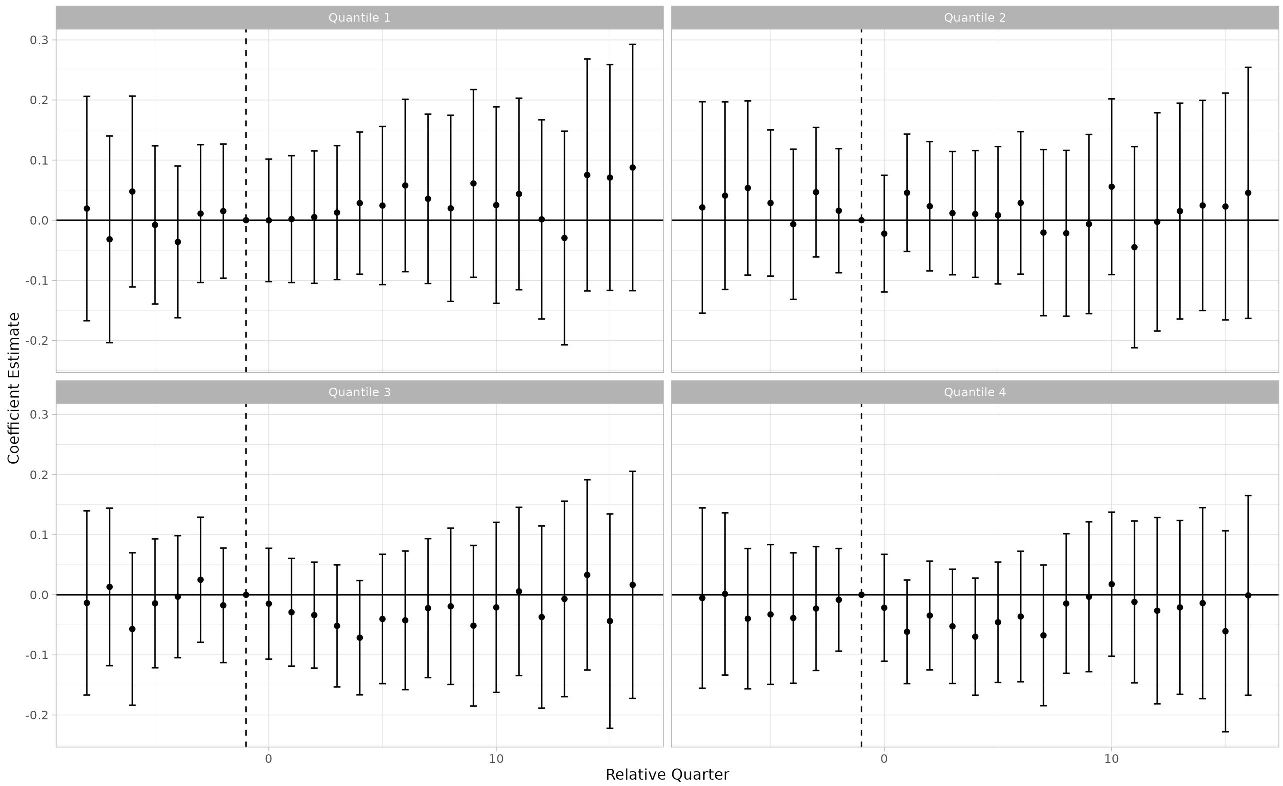


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level. Income Rank 1 is the lowest income, Rank 4 is the highest income.

Figure A.27: Log Total Expenditure for Non-Urban and Urban

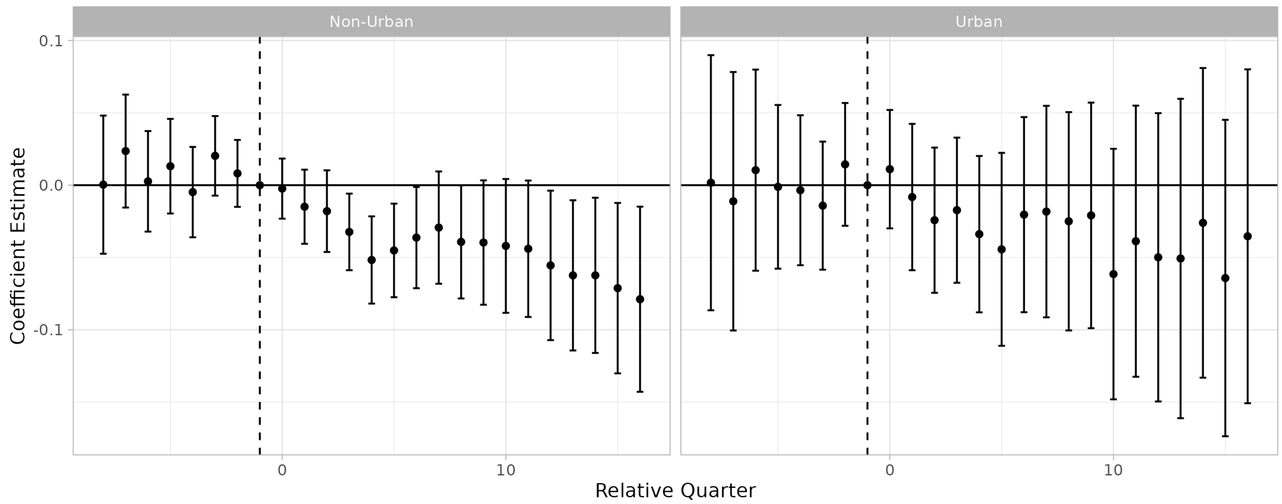


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.28: Food Expenditure at the Dollar Store for Non-Urban and Urban

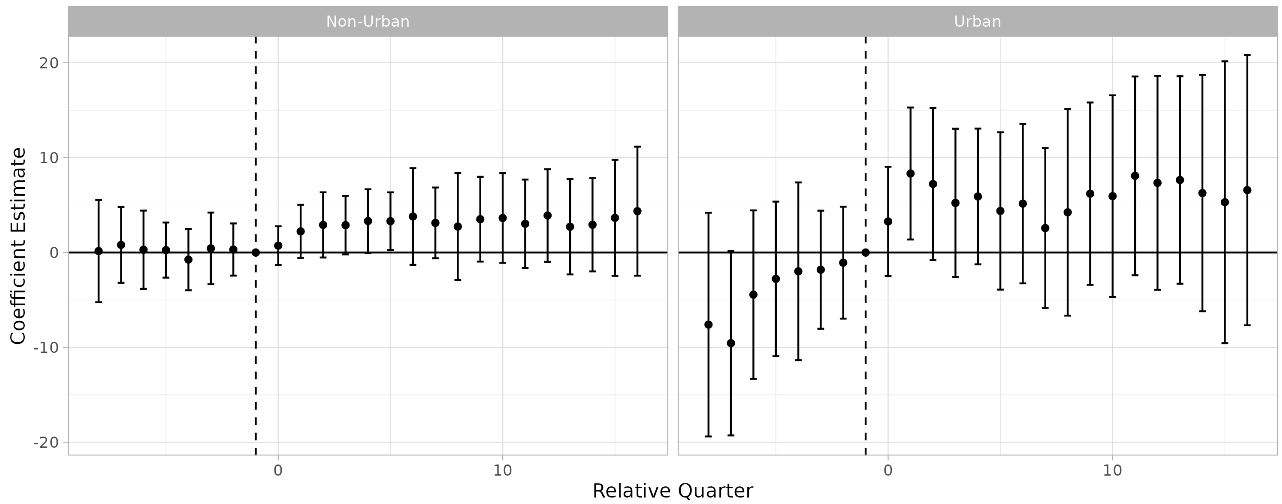


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.29: Food Expenditure at the Grocery Store for Non-Urban and Urban

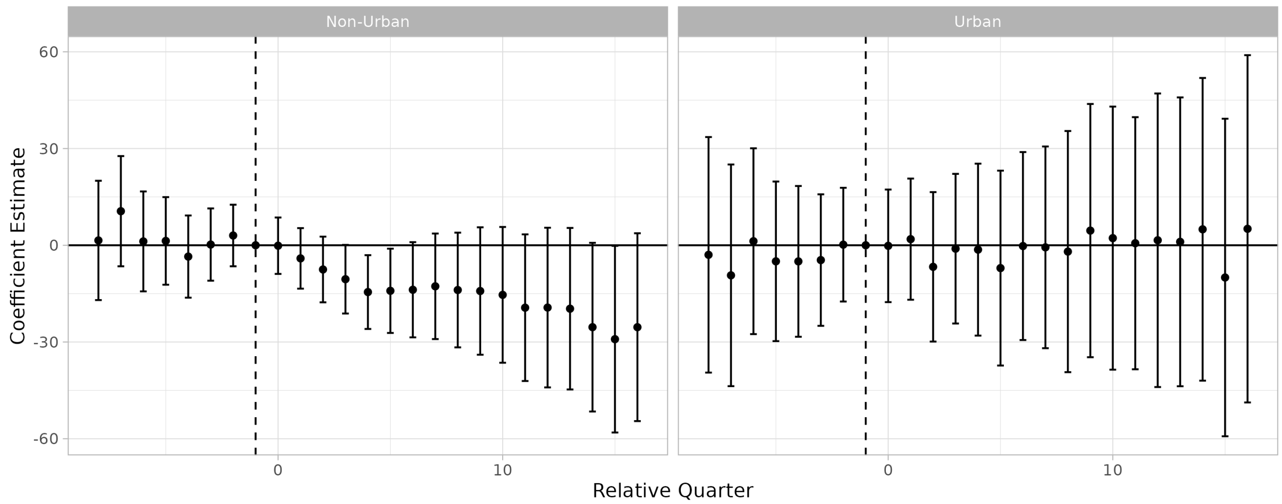


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.30: Log Relative Price Index for Non-Urban and Urban

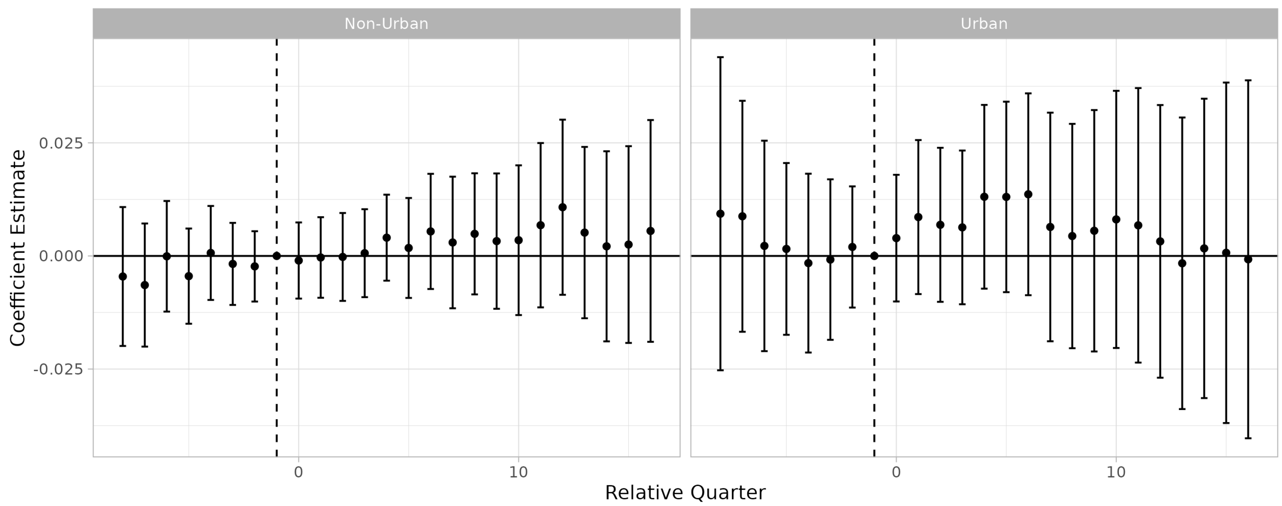


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.31: Log Number of Unique Varieties for Non-Urban and Urban

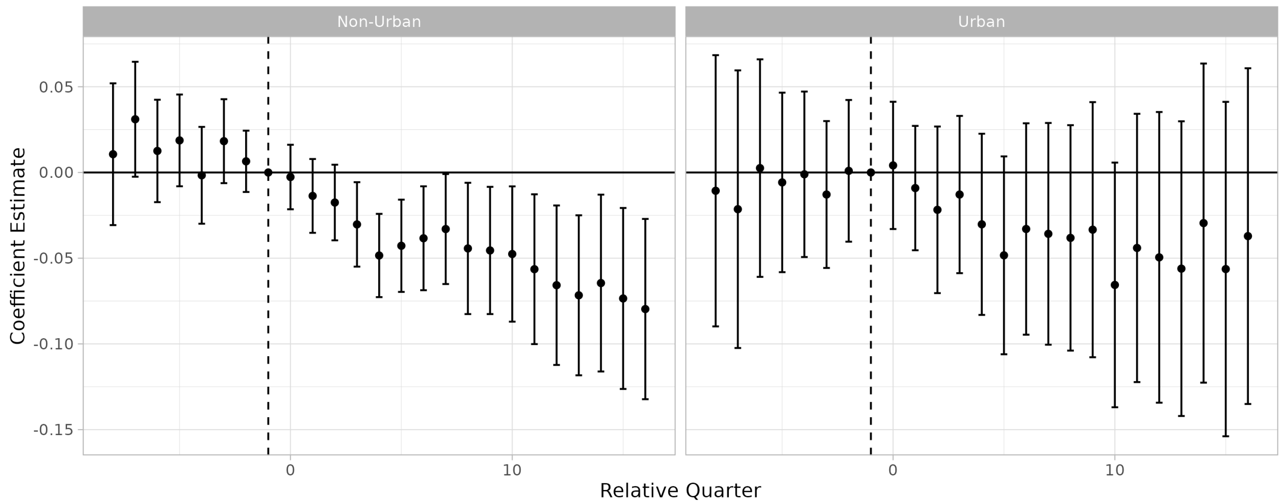


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.32: Log Quantity of Fresh Produce for Non-Urban and Urban

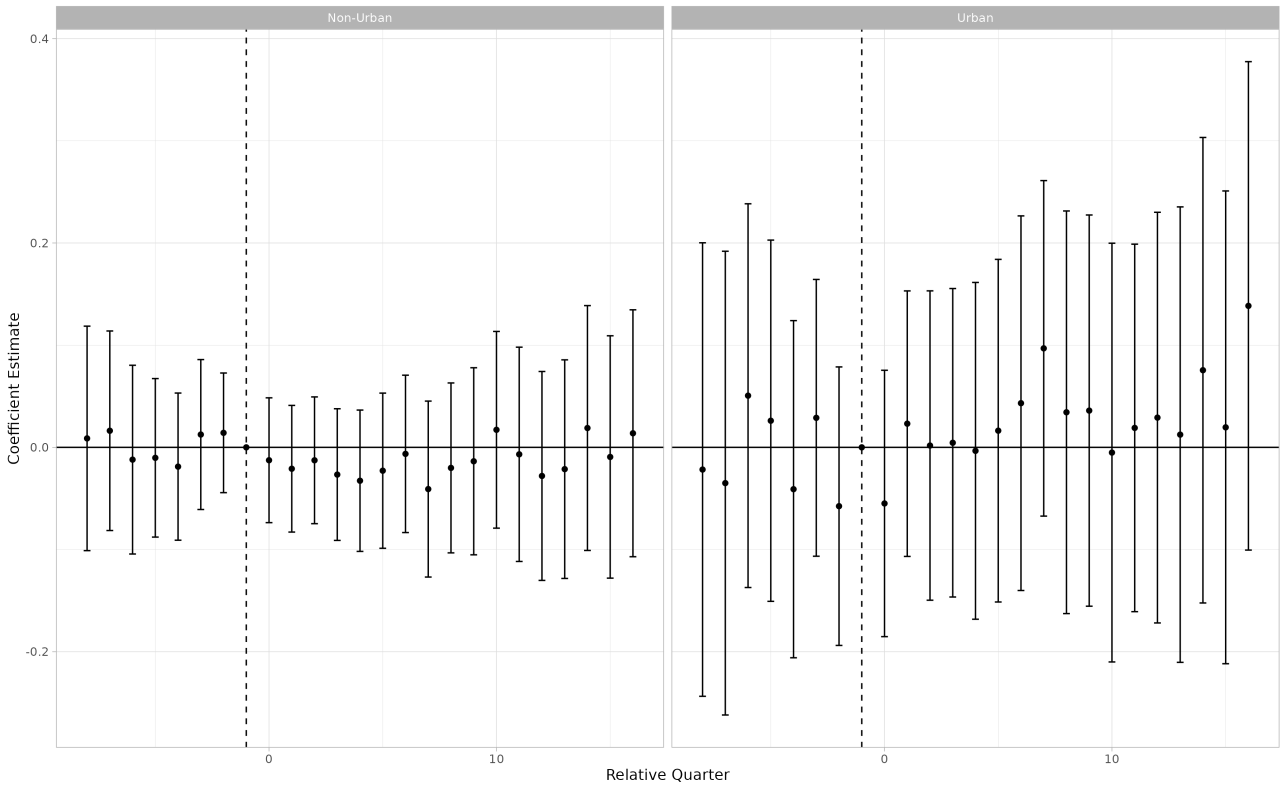


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.33: Log Total Expenditure for Food Deserts

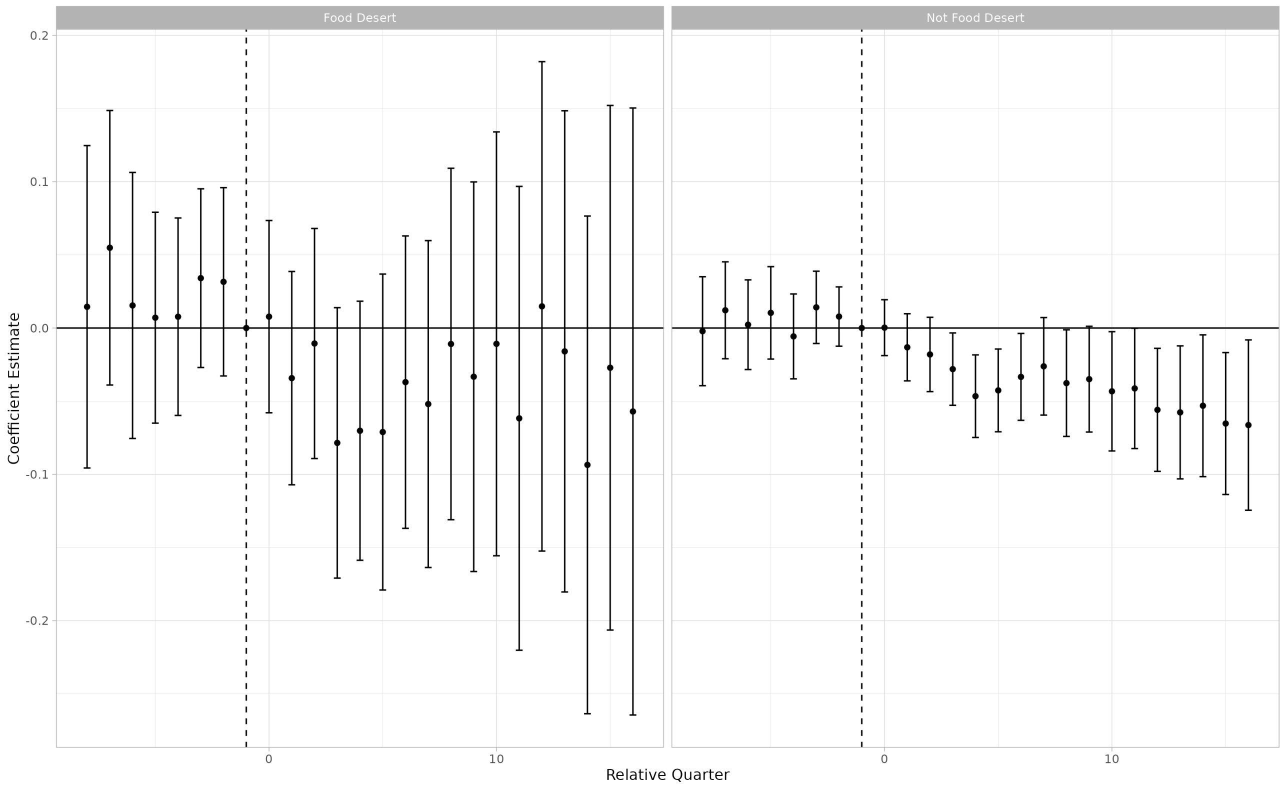


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.34: Food Expenditure at the Dollar Store for Food Deserts

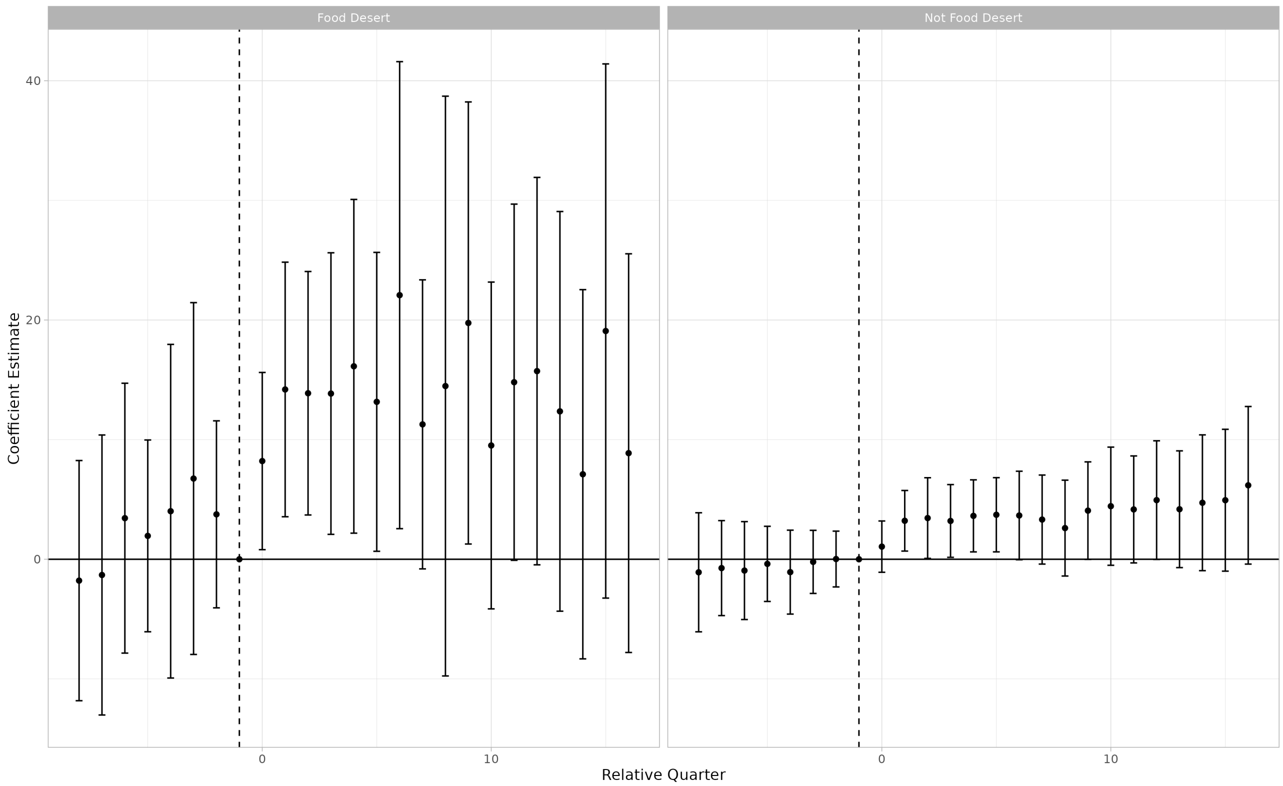


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.35: Log Relative Price Index for Food Deserts

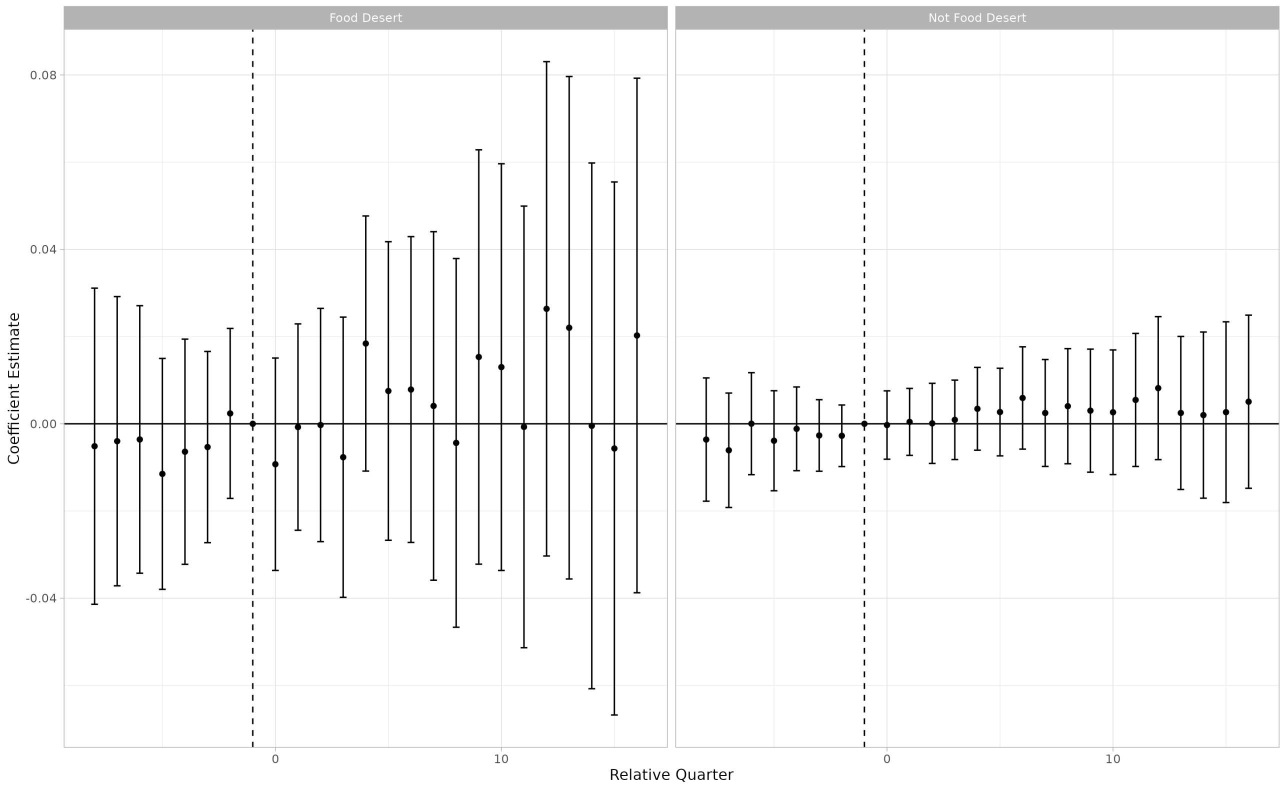


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.36: Log Number of Unique Varieties for Food Deserts

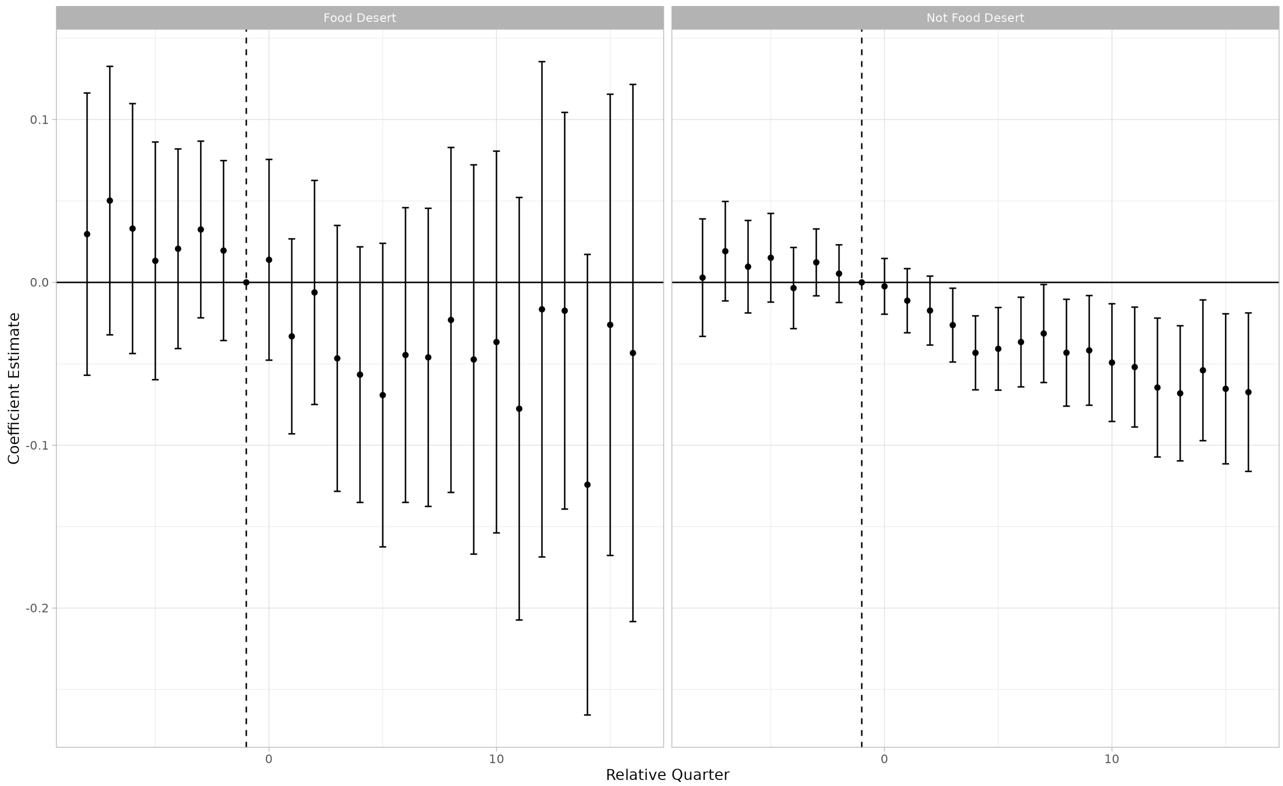


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

Figure A.37: Log Quantity of Fresh Produce for Food Deserts

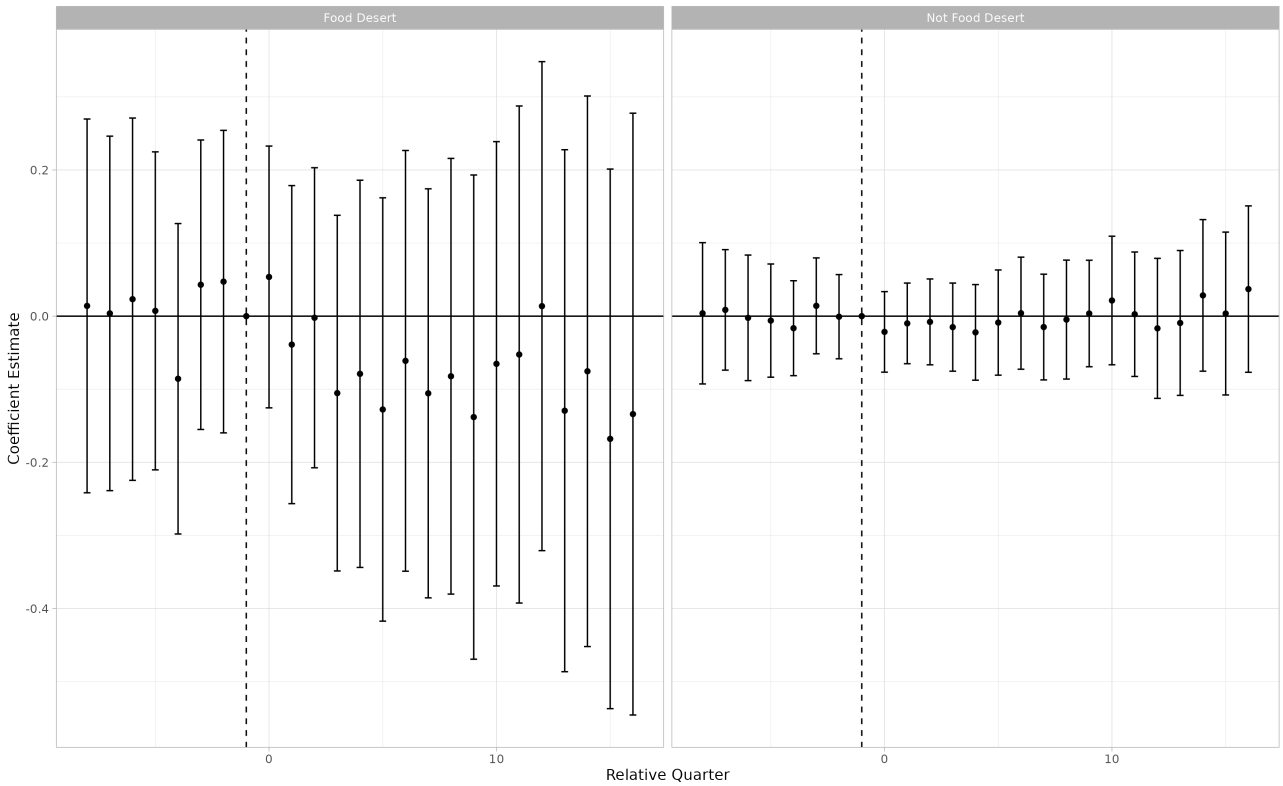


Figure reports event study estimates with 95% confidence intervals from Equation 5, using 2008-2019 Homescan data. The analysis uses a heterogeneity-robust estimator proposed by Callaway and Sant'Anna (2021). Observations are not weighted for national representativeness. Errors are clustered at the zipcode level.

A.3 Demand Estimation: Disentangling Price, Size, Variety, and Convenience

In this section, we employ a slightly modified demand specification to understand the role of convenience in shopping at the dollar store. Mechanically, dollar store entry increases the set of stores available for a household to shop at, and dollar store customers may benefit from additional convenience. Estimating the value of convenience is important as the value of convenience may matter for welfare and policy; for example, if the household values the dollar store because of its convenience, then the policy recommendation – holding all else fixed – is to increase ease of access to dollar stores and thus encourage further dollar store entry. However, estimating the value of convenience is challenging because convenience itself is not measurable.

Since we do not know where people live, we use number of trips to proxy for convenience (a less precise alternative would be to use distance between the dollar store and the centroid of the zip code). This revealed-preference style approach assumes that households will shop more at stores that are more convenient, and shop at less at stores that are less convenient. As a result, we include the number of trips made to a store type, $d_{s(j)}$. We note that trips are an outcome variable and not an input to preferences, and for this reason, we include number of trips only as robustness to the main specification.

$$u_{ij} = -\underbrace{\alpha \log p_j}_{price} + \underbrace{\beta v_j}_{variety} + \beta_2 v_j^2 + \underbrace{\eta m_j}_{product\ size} + \eta_2 m_j^2 + \underbrace{\psi_{g(j)}}_{product\ group} - \underbrace{\gamma d_{s(j)}}_{\text{trips}} + \epsilon_{ij}(\lambda)$$

We report the results in Table A.11 and A.12. We find that our coefficients on prices, size, variety, and the nesting parameter are very similar to the baseline specification in Table 3. As for convenience, we find a negative and significant coefficient on non-grocery trips, indicating that trips to retailers that are not the grocery store are inconvenient and provide disutility. This dovetails with the Allcott et al. (2019) results that the grocery store is a necessary good, and the other store types are much less necessary. We note that club stores are likely negative because one tends to buy in bulk at the club store (and possibly also the discount store), and so probably does not take many trips there. Meanwhile, the grocery, dollar store, convenience store, and drug stores are subject to making many trips, and so are likely reflective of preferences of these stores types.

We re-compute welfare under this alternative specification which includes trips to each store type and we report the results in Table A.13. We find that the benefit to the average household range from 4.7% for the average household in the top income rank to 11% for the average household in the bottom income rank. The average household is in rank 2 and spends 598$ per year on food expenditures, which translates into a 47$ value of the dollar store to the average household.

A.4 Retail Closing Instrument in the Context of Dollar Store Entry

We illustrate the instrument in the context of dollar store entry, exploiting the fact that dollar stores only move into stores of particular sizes (8,000-10,000 square feet). We show that the dollar store enters into zip codes previously populated by now-bankrupt stores if the bankrupt retailers also tended to occupy storefronts between 8,000 and 10,000 square feet, but that the dollar store is less likely to enter into zip codes previously populated by now-bankrupt stores with store sizes smaller than 8,000 square feet or larger than 10,000 square feet.

First, we present anecdotal evidence which shows that dollar stores enter locations of approximately 8,000-10,000 square feet:

> "If you want to be profitable, start with an 8,000-square-foot store." – Wally Lee, director of marketing and technology of supplier of dollar stores.[38]

> "Our stores predominantly range from 8,000 - 10,000 selling square feet" – Dollar Tree Annual Report, 2020

> "We lease the vast majority of our stores ... this leasing strategy [allows us] to pursue various expansion opportunities resulting from changing market conditions" – Dollar Tree Annual Report, 2020

Next, we show that the first dollar store is more likely to enter a zip code following the bankruptcy of a storefront that typically rented out a similar-sized store. In the first table, we use include stores with the "correct" size for dollar store entry (stores between 8,000 and 10,000 square feet), and in the second, we include stores with the "incorrect" size for dollar store entry (stores less than 8,000 square feet or larger than 10,000 square feet).

We include a different specification than the demand estimation. That is, in this regression the specification is now

$$D_{it} = \alpha + \sum_k \beta_k Z_{ik} + \sigma_i + \lambda_t + \epsilon_{it} \tag{13}$$

where D_{it} is a binary variable for whether the first dollar store has already entered zip code i by time t, Z_{ik} as a binary variable equal to one if a retailer went bankrupt in time period $t-k$ operated a store in zip code i at that time, σ_i and λ_t are zip and time fixed effects. We expect the instrument to explain much less of the variation in dollar store entry because there are retailers that go bankrupt after the first dollar store entry, that in no way could induce the first dollar store to enter that zip code. Results show that dollar store entry is positively correlated with correctly-sized bankruptcy retailer exit (Table A.8), and negatively correlated with incorrectly-sized bankruptcy retailer exit (Table A.9).

[38] Excerpt from New York Times article "The Dollar-Store Economy"

References

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

Ailawadi, Kusum L., Jie Zhang, Aradhna Krishna, and Michael W. Kruger, "When Wal-Mart Enters: How Incumbent Retailers React and how this Affects their Sales Outcomes," *Journal of Marketing Research*, August 2010, *47* (4), 577–593.

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell, "Food Deserts and the Causes of Nutritional Inequality*," *The Quarterly Journal of Economics*, November 2019, *134* (4), 1793–1844.

Arcidiacono, Peter, Patrick Bayer, Jason R. Blevins, and Paul B. Ellickson, "Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition," *The Review of Economic Studies*, July 2016, *83* (3), 889–931.

_ , Paul B Ellickson, Carl F Mela, and John D Singleton, "The Competitive Effects of Entry: Evidence from Supercenter Expansion," May 2019, p. 56.

Atkin, David, Benjamin Faber, and Marco Gonzalez-Navarro, "Retail Globalization and Household Welfare: Evidence from Mexico," *Journal of Political Economy*, February 2018, *126* (1), 1–73.

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

Basker, Emek and Michael Noel, "The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry," *Journal of Economics & Management Strategy*, December 2009, *18* (4), 977–1009.

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

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

Cao, Yue, "The Welfare Impact of Dollar Stores," November 2022, p. 61.

Caoui, El Hadi, Brett Hollenbeck, and Matthew Osborne, "The Impact of Dollar Store Expansion on Local Market Structure and Food Access," *SSRN Electronic Journal*, 2022.

Chenarides, Lauren, Metin Cakir, and Timonthy J. Richards, "Dollar Store Entry," November 2021.

Conlon, Christopher and Jeff Gortmaker, "Best practices for differentiated products demand estimation with PyBLP," *The RAND Journal of Economics*, 2020, *51* (4), 1108–1161.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform Pricing in U.S. Retail Chains*," *The Quarterly Journal of Economics*, 06 2019, *134* (4), 2011–2084.

Ellickson, Paul B. and Paul L.E. Grieco, "Wal-Mart and the geography of grocery retailing," *Journal of Urban Economics*, May 2013, *75*, 1–14.

Feng, Wenhui, Elina T. Page, and Sean B. Cash, "Dollar Stores and Food Access for Rural Households in the United States, 2008–2020," *American Journal of Public Health*, 2023, *113* (3), 331–336. PMID: 36657091.

Handbury, Jessie, "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities," *Econometrica*, 2021, *89* (6), 2679–2715.

Hausman, Jerry and Ephraim Leibtag, "Consumer benefits from increased competition in shopping outlets: Measuring the effect of Wal-Mart," *Journal of Applied Econometrics*, December 2007, *22* (7), 1157–1177.

_ , Gregory Leonard, and J. Douglas Zona, "Competitive Analysis with Differenciated Products," *Annales d'Économie et de Statistique*, 1994, (34), 159–180.

Horowitz, Joel L., "Chapter 52 - The Bootstrap," 2001, *5*, 3159–3228.

Illanes, Gastón and Sarah Moshary, "Market Structure and Product Assortment: Evidence from a Natural Experiment in Liquor Licensure," Technical Report w27016, National Bureau of Economic Research, Cambridge, MA April 2020.

Jia, Panle, "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 2008, *76* (6), 1263–1316.

Leung, Justin and Zhonglin Li, "Big-Box Store Expansion and Consumer Welfare," *SSRN Electronic Journal*, 2021.

Natan, Olivia R, "Choice Frictions in Large Assortments," 2020.

Schmall, Alexandria, Rachel Kinney, Alexa Gozdiff, Madison Reid, Sara Folta, Wenhui Feng, and Sean Cash, "Food Shopping Experiences Among Dollar Store Shoppers in Fresno and Boston During the COVID-19 Pandemic," *Current Developments in Nutrition*, June 2021, *5* (Supplement_2), 246–246.

Shannon, Jerry, "Dollar Stores, Retailer Redlining, and the Metropolitan Geographies of Precarious Consumption," *Annals of the American Association of Geographers*, June 2021, *111* (4), 1200–1218.

Verboven, Frank, "The nested logit model and representative consumer theory," *Economics Letters*, 1996, *50* (1), 57–63.