ESSAYS ON CONSUMER BEHAVIOR AND RETAIL REGULATION

Ethan Mallick Hossain

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

____

Katja Seim
Professor of Economics

Graduate Group Chairperson

____

Jesús Fernández-Villaverde
Professor of Economics

Dissertation Committee

Holger Sieg, Professor of Economics

Aviv Nevo, Professor of Economics

Sarah Moshary, Assistant Professor of Marketing

ESSAYS ON CONSUMER BEHAVIOR AND RETAIL REGULATION

For my parents, my sister, and my wife, who taught me that the world is filled with wonders, if I look and listen carefully.

ACKNOWLEDGMENT

I am indebted to my advisor, Katja Seim, as well as the other members of my committee, Holger Sieg, Aviv Nevo, and Sarah Moshary, for your invaluable guidance, patience, and encouragement. Without your help, I would have been unable to write this dissertation. In particular, I want to thank Katja Seim for spending countless hours making sure I continued moving forward and encouraging me whenever I stumbled along the rocky research road. Thanks to my coauthor, Jason Sockin, for never losing sight of the bigger picture. I hope that I have taught you a fraction of what you have taught me. Thank you also to Claudia Sahm, Emek Basker, Frank DiTraglia, Petra Todd, Andrew Shephard, and Francesco Agostinelli for helping me make my research accessible to a wide array of audiences by providing extensive feedback on my writing and presentations.

I also want to thank my grad school friends: Minsu Chang, Amanda Chuan, Assa Cohen, María José Orraca Corona, Stuart Craig, Matt Davis, Emma Dean, Ryan Fackler, Caitlin Gorback, João Granja, Pengfei Han, Kathleen Hui, Dan Kaliski, Ashwin Kambhampati, Nitin Krishnan, Tomás Larroucau, Joonbae Lee, André Victor Luduvice, Paolo Martellini, Stefano Pietrosanti, Isaac Rabbani, Eugenio Rojas, Paul Sangrey, Takeaki Sunada, Gabrielle Vasey, and Hanna Wang. Thanks for your friendship, support, and advice throughout this journey. I could not have done it without you. The future of economics is bright because of you and I am thankful to have benefited from the unique perspective each of you has shared with me.

Last, but certainly not least, I must thank my family. To Mom and Dad for giving me the bedrock foundation to do anything and never telling me something was impossible until after I had done it. To my sister Shyla for her unwavering love and support, always knowing when I needed something before I knew I needed it. To my wife Noora for always being by my side even when we were hundreds of miles apart. I cannot imagine taking this journey without each of you.

ABSTRACT

ESSAYS ON CONSUMER BEHAVIOR AND RETAIL REGULATION

Ethan Mallick Hossain

Katja Seim

The retail sector is one of constant innovation, where retailers relentlessly compete for customers by providing new products, better service, and attractive prices. At the same time, regulators must keep pace with these innovations to ensure that markets are fair and consumers are protected. This dissertation is composed of three chapters that examine the interactions of consumer behavior and retail regulation.

The first chapter examines how bulk buying varies by household income and analyzes the factors that affect a household's bulk buying decision. Using reduced-form methods and detailed household-level purchase data, I show that many factors, including the cognitive costs of computing unit prices, store preferences, storage costs, and budget constraints, affect a household's bulk buying. I then estimate a discrete-choice model that incorporates cognitive costs and storage costs and find that mandating the display of unit prices would substantially increase bulk buying and lower the unit prices paid by households, especially low-income households.

The second chapter studies how imposing sales taxes on previously tax-free online purchases affects household shopping behavior. Historically, e-commerce was an easy way for consumers to avoid sales taxes, but over the past decade, online retailers were required to collect sales taxes, negating the structural price advantage they had. Using detailed online shopping and browsing data, I find that in response to sales tax collection, households reduce their spending at taxed online retailers, but find no evidence that households change their search behavior or offline shopping expenditures.

The third chapter analyzes whether welfare transfers are linked to lottery gambling. A minority of lottery retailers are eligible to accept Supplemental Nutrition Assistance Program (SNAP) benefits, but these stores account for a majority of lottery sales. By combining novel data on store-level lottery sales with a range of policy shocks to the SNAP program, this chapter finds that SNAP benefits decrease lottery gambling,

likely by decreasing shopping frequency, and therefore, the number of lottery gambling opportunities.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Chapter 1

LESS IS MORE EXPENSIVE: INCOME DIFFERENCES IN BULK BUYING

By E. Mallick Hossain

1.1 Introduction

Grocery purchases account for a sizable share of a household's discretionary spending, especially for the lowest-income households (BLS 2019). To save money, households often wait for sales, redeem coupons, or purchase generic brands (Griffith et al. 2009). Low-income households also increase their home production to reduce their spending (Aguiar and Hurst 2005). Quantity discounts are another way for households to save money. Even though households buy more on a given shopping trip, they pay lower unit prices, and reduce their overall spending.

My paper examines how large these quantity discounts are, how bulk buying varies by income groups, and which factors influence the decision to buy in bulk. Most explanations for why households pay different prices for the same product relate to differences in search behavior. Prices vary based on where households shop or whether they use coupons (Talukdar 2008; Griffith et al. 2009). However, even in the absence of sales and coupons, prices for the same product can differ within a particular store due to quantity discounts.

This paper contributes new findings that, despite the substantial savings available from quantity discounts, low-income households are less likely to buy in bulk than high-income households.[1] Kunreuther (1973) provides the first evidence of this "bulk buying gap" for a few specific products and Orhun and Palazzolo (2019) expands this finding to a whole product category. Since households purchase a variety of products when shopping, I show that the bulk buying gap exists across the full range of product categories that households purchase.

I find that cognitive costs, store preferences, budget constraints, and storage costs all contribute to this bulk buying gap. First, the cognitive costs of assessing price differences across products can prevent households from making economical decisions (Mitchell, Lennard, and McGoldrick 2003). Providing price information reduces the effort needed to compare prices, and households change their purchase decisions when relevant price information is displayed prominently (Chetty, Looney, and Kroft 2009; Bogomolova and Jarratt 2016). Posting unit prices reduces the cognitive costs of comparing unit prices across different products. Sixteen states have regulations governing the display of unit prices, but no study has evaluated the impact of these regulations on consumer behavior. I provide the first nationwide study of the impact of displaying unit prices on bulk purchasing and find that households are significantly more likely to buy in bulk when retailers are mandated to display unit prices.

Second, households may also make different purchase decisions based on where they live or where they choose to shop (Chung and Myers Jr 1999; Talukdar 2008; Allcott et al. 2019). I show that even within neighborhoods, there are large differences in bulk buying between high- and low-income households. On the other hand, income differences in bulk buying are attenuated, but still substantial, after conditioning on the type of store where households choose to shop.

Third, budget constraints affect bulk buying because low-income households may not have enough cash on hand to purchase a bulk package. Leveraging within-month variation in budgets, I show that the lowest-income households slightly decrease their bulk purchasing towards the end of month, presumably when their budget constraint is binding. In contrast, middle- and high-income households either do not change their bulk buying or slightly increase their bulk buying towards the end of the month.

Fourth, storage costs also affect the bulk buying decision because even though large packages provide lower unit prices, they are more cumbersome to store. I

[1] Throughout this paper, "high-income" refers to households making over $100,000 and "low-income" refers to households making under $25,000.

show that households are more likely to buy in bulk when they live in larger homes and when products are smaller. I provide a new approach to estimate storage costs cross-sectionally using differences in product "concentration." One way of identifying household-level storage costs compares purchase frequencies of households with the same demand for a product (Hendel and Nevo 2006). A household with high storage costs will purchase small packages more frequently. I propose a complementary approach: I compare purchases of products that are otherwise identical, but differ in their level of concentration. All else equal, households who buy smaller, more concentrated packages have higher storage costs than those that buy larger, less concentrated packages. Based on this approach, I find that low-income households have higher storage costs than high-income households. On balance, this means that firms extract higher rents from low-income households because they have less ability to store for future consumption.

For my analysis, I combine household- and store-level datasets to study income heterogeneity in bulk buying. Nielsen's Consumer Panel data is a nationally representative panel survey of household grocery purchases, and Nielsen's Retail Scanner data is a national panel of weekly UPC-level sales data from over 30,000 stores. I construct a new dataset of state-level per-unit pricing regulations, including a measure of regulatory stringency. I also use data on entry dates and locations of over 1,400 warehouse clubs in the United States. As a result, I have a comprehensive view of a household's possible product choices, available price information, retail environment, and resulting expenditures.

I find that low-income households could realize substantial savings from buying in bulk at the same rate as high-income households. To do this, I estimate the average bulk discount for each product category based on Nielsen's weekly store-level price and product data. The average discount across all product categories is such that a 10% larger package has a 5% lower unit price. Then, I estimate how much each household buys in bulk using Nielsen's household-level purchase data. Given each product category-specific bulk discount and household-level bulk buying, I predict how much low-income households could save if they increased their bulk buying intensity to match that of high-income households. I find that low-income households would reduce their annual grocery expenditures by 5% if they bought in bulk like high-income households, saving an aggregate of \$5.4 billion annually.

I then employ three differences-in-differences models to determine how much cognitive costs, store preferences, and budget constraints affect bulk buying. The first model uses a novel dataset that I compiled of state regulations regarding the display of per-unit prices and exploits the fact that a significant share of households in the Nielsen Consumer Panel moves between regulatory regimes when they move from one state to another. Before households move to a state without unit price posting requirements, their bulk buying behavior is similar to that of households that remain in the same regulatory regime. After households move, however, I find that their bulk buying is 4–5% lower than that of households who did not experience a regulatory change.

The second differences-in-differences model shows how much store preferences, particularly for warehouse clubs, affect bulk buying. The bulk buying gap narrows substantially after controlling for the types of stores households shop at; high-income households spend a larger portion of their budget at warehouse clubs. To estimate the effect of warehouse clubs on bulk buying, I examine how bulk buying changes within households before and after a warehouse club enters nearby. Before a warehouse club enters, bulk buying is similar for households that will and will not experience a warehouse club entry. After a warehouse club enters, however, I find bulk buying increases by 5–10% compared to households that did not experience an entry within a 15-mile radius. This increase is only limited to middle- and high-income households and is due to increased shopping at warehouse clubs.

The third differences-in-differences model analyzes how budget constraints affect bulk buying. To do this, I analyze within-household changes in bulk buying over the course of a month because the liquidity of poorer households decreases towards the end of the month (Orhun and Palazzolo 2019). Using information on weekly bulk buying from Nielsen, I find that low-income households slightly decrease their bulk buying by 0.5–1% between the first half and second half of the month while middle- and high-income households slightly increase their bulk buying by 0.5–1.5% over the same period.

I also assess the importance of storage costs using differences in bulk buying relative to the size of a household's home and relative to the physical size of the product being purchased. I examine how a household's bulk buying changes when it moves to a different type of housing, after controlling for other within-household changes. Bulk buying is 3–4% higher for the same household when it lives in a single-family home relative to when it lives in an apartment, controlling for demographic differences between households. I also find that the bulk buying gap is smaller for products with smaller physical footprints.

Finally, I construct a discrete-choice model of consumer purchasing behavior to quantify consumer preferences and disentangle the contributions of cognitive and storage costs to the bulk buying decision. I estimate this model using data on toilet paper purchases. Households choose a product based on price, quantity, quality, and package size, which serves as a proxy for storage costs. I can separate preferences for quantity from size preferences because I demonstrate that toilet paper comes in varying "concentrations." I allow state-level unit pricing mandates to affect a household's unit price sensitivity. From this demand model, I simulate household responses to two counterfactuals: 1) universally posting unit prices and 2) reducing storage costs.

My model predicts that requiring stores to post unit prices would reduce the bulk buying gap in package size purchased between high- and low-income households by 26%. Reducing storage costs would close the remainder of the gap and low-income households would actually buy more in bulk than high-income households. As a result of these policies, households would buy larger quantities of toilet paper and pay lower unit prices. Universally displaying unit prices would encourage households to better utilize quantity discounts by reducing cognitive costs, increasing bulk buying, and helping consumers save money.

The rest of the paper is structured as follows. Section 1.2 describes the data. Section 1.3 documents new facts of quantity discounting. Section 1.4 presents evidence of contributing factors to the bulk-buying gap. Section 1.5 introduces the model. Section 1.6 presents estimation results. Section 1.7 details the counterfactual exercises and Section 1.8 concludes.

1.2 Data

In this section, I describe the datasets used for my analysis and give a brief overview of their respective features.[2] Nielsen's Consumer Panel data provides information on households' shopping and purchasing decisions. Nielsen's Retail Scanner data provides information on weekly product assortments and prices. A new regulatory dataset I construct contains information on state-level regulations regarding the display of per-unit pricing. Finally, warehouse club data provides information on the location and entry dates of warehouse clubs across the United States. By combining these data, I have a comprehensive view into a household's possible product choices, available price information, retail environment, and their resulting purchase decision.

Nielsen Consumer Panel Data

I use the Nielsen Consumer Panel dataset from 2004–2017. This dataset is a panel of about 178,000 unique households. I observe about 40,000 households each year from 2004–2006 and about 60,000 households each year from 2007–2017. Households scan all items that they purchase and then input information about quantities, prices, date of purchase, and store. Nielsen retains about 80% of its panel from year to year with the mean and median tenure of a household being four and three years, respectively.

I consider food, drink, and non-food grocery (e.g., paper towels, toilet paper, detergent, etc.) purchases made at grocery stores, discount stores, dollar stores, warehouse clubs, and drug stores. These outlets account for over 90% of household expenditures in these categories. I exclude alcohol, tobacco, health, and general merchandise products from my analysis since these products (e.g., cigarettes, painkillers, etc.) may have different consumption patterns than grocery products or are not suited for bulk purchases (e.g., printers, cookware, linens). I also exclude households with a student or military head of household as well as those with an annual income of less than $5,000 and those living in mobile homes. Only about 7% of households are excluded and I use the remaining 166,000 households for my analysis. See Appendix 1.A for further details of sample construction.

[2] Researcher's own analyses derived based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Nielsen computes projection weights to ensure their sample is nationally representative. Weights are calculated to match population moments based on household size, income, age, race, ethnicity, education, occupation, and presence of children. All analyses use these projection weights unless otherwise stated. Nielsen groups household income into 16 different income bins. Due to the large number of bins, in tables and parts of the text, I will report differences by income quartiles. However, where possible (especially in graphs), I will report estimates for each income bin. Table 1.1 presents descriptive statistics for households in the sample.

Table 1.1: Nielsen Consumer Panel Summary Statistics

Variable	Mean	SD	25th Pctile	75th Pctile
Household income ($000s)	57.13	31.17	27.5	85
Household size	2.56	1.45	1	3
Age	52.59	14.39	41.5	63
College educated	0.39	0.49	0	1
Child present	0.33	0.47	0	1
Married	0.50	0.50	0	1
N (Household-Years)			734,724	
N (Households)			166,163	

Note: Household income is grouped into bins. Midpoints of each bin are used in order to calculated sample moments. Data are weighted for national representativeness.

Nielsen Scanner Data

The Nielsen Scanner data contains average weekly prices and volume sold of individual products at about 35,000 stores from about 90 retail chains between 2006–2016. Average prices are weighted by the volume sold. Only products with positive sales in a given week are recorded. I match the Retail Scanner data with the Consumer Panel data based on store identification numbers and purchase dates. By matching the two datasets, I recover the set of products available to a household and the product it chose to purchase.

Unit Pricing Regulations

I compile a novel dataset on state-level regulations regarding the display of unit prices. The data is based on annual regulatory updates aggregated in Handbook 130 published by the National Institute of Standards and Technology (NIST 2019). I cross-check this information with state regulatory codes and state officials to ensure accuracy. This data includes information on which states have regulations, when they were adopted, and how stringent these regulations are. More details are discussed in Section 1.4.

Warehouse Club Data

I also use data on all warehouse clubs in the United States between 2004–2015 gathered by Coibion, Gorodnichenko, and Koustas (2017). This data records information on the opening dates, locations, and identity of all warehouse clubs in the United States. It was gathered by combining information available on company websites, annual reports, and by contacting firms.

1.3 Stylized Facts

In this section, I document two new facts about quantity discounts. First, I show that quantity discounts apply to 91% of grocery categories. Second, I document that households making over $100,000 are 26% more likely to buy non-food items in bulk than households making $5,000–$8,000 annually, compared to only 3% for food items. Combining these findings, I estimate that low-income households could reduce their grocery expenditures by 5%, saving an aggregate of $5.4 billion annually, simply by buying in bulk at the same rate as high-income households.

Quantity Discount Prevalence

Quantity discounts are a specific form of non-linear pricing in which unit prices decrease as package size increases. To establish the prevalence and magnitude of quantity discounts, I use Nielsen's Retail Scanner data from 2016. I estimate quantity discounts using the following regression for each of the 693 product categories:[3]

$$\ln(P)_{ibm} = \alpha + \beta \ln(Size)_{ibm} + \lambda_{bm} + \epsilon_{ibm}, \tag{1.1}$$

where P_{ibm} is the unit price (package price divided by package size) of product i from brand b purchased in market m (defined as a store-week). $Size_{ibm}$ is the item's package size, which is the number of units included in a UPC (e.g., quart, square feet, count, pound, etc.). λ_{bm} is a brand-store-week fixed effect. Variation in unit prices across package sizes within the same brand-store-week identify β.[4] If retailers offer quantity discounts, then β will be negative.

Figure 1.1 plots the distribution of β across product categories (statistically insignificant betas are zero). 91% of all product categories have a statistically significant and negative β and non-food items generally have larger discounts than food items.[5] The median β is -0.51 for non-food products, which means that a 10% increase in package size is associated with a 5.1% decrease in unit price. This discount is larger than the median β for food items (-0.43).[6] The size and near-universality of quantity discounts suggest they offer substantial savings to households without sacrificing consumption.[7]

Bulk Purchasing

Given how common and how large quantity discounts are, households can use quantity discounts to save money on a wide range of items. However, since food products deteriorate while non-food products do not, bulk buying will likely differ between

[3] 48 categories could not be estimated typically because the data did not have sufficient variation. These were generally uncommon categories like mushroom sauce, canned grapes, and canned chow mein.

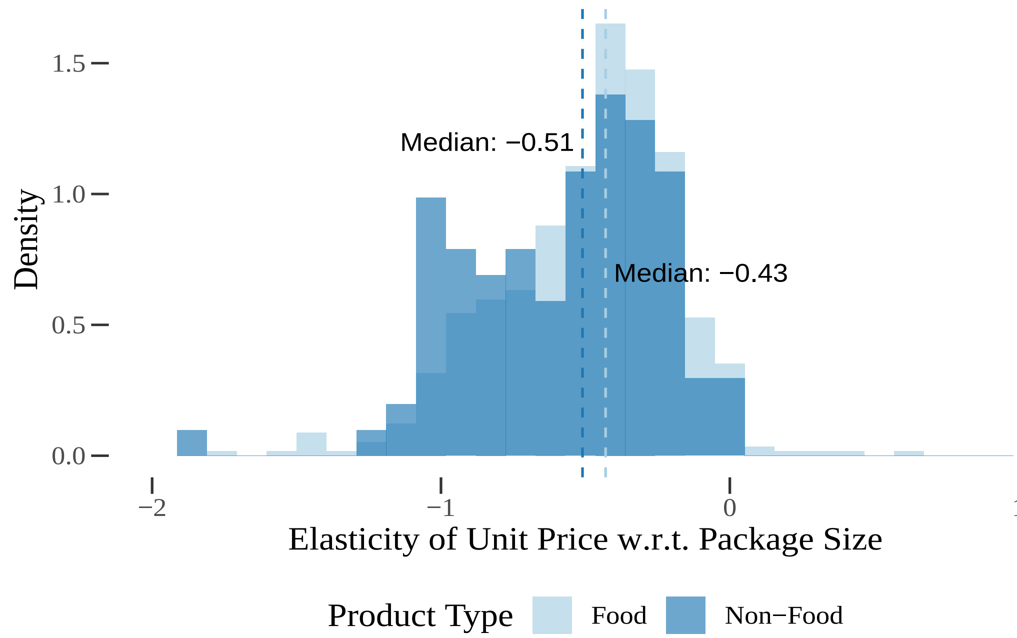
[4] Some readers may be concerned that the positive sales threshold limits the number of weeks products are observed. I find that a large majority of products (at the UPC level) are observed for over half of the year. The unobserved weeks can be attributed to a variety of reasons including zero sales, discontinued products, or missing reports from retailers. Observing products for most weeks of the year limits the possibility that quantity discounts are estimated on a limited subset of weeks.

[5] Some products do have a significantly positive coefficient, indicating that unit prices increase with package size. These quantity "surcharges" are less common, but have been highlighted before (Sprott, Manning, and Miyazaki 2003).

[6] These findings are robust to outliers. Winsorizing unit prices at the 98th and 90th percentile produces almost identical estimates.

[7] For a comparison of quantity discounts with coupons, see Appendix 1.A.

Figure 1.1: Distribution of Bulk Discounts by Product Type



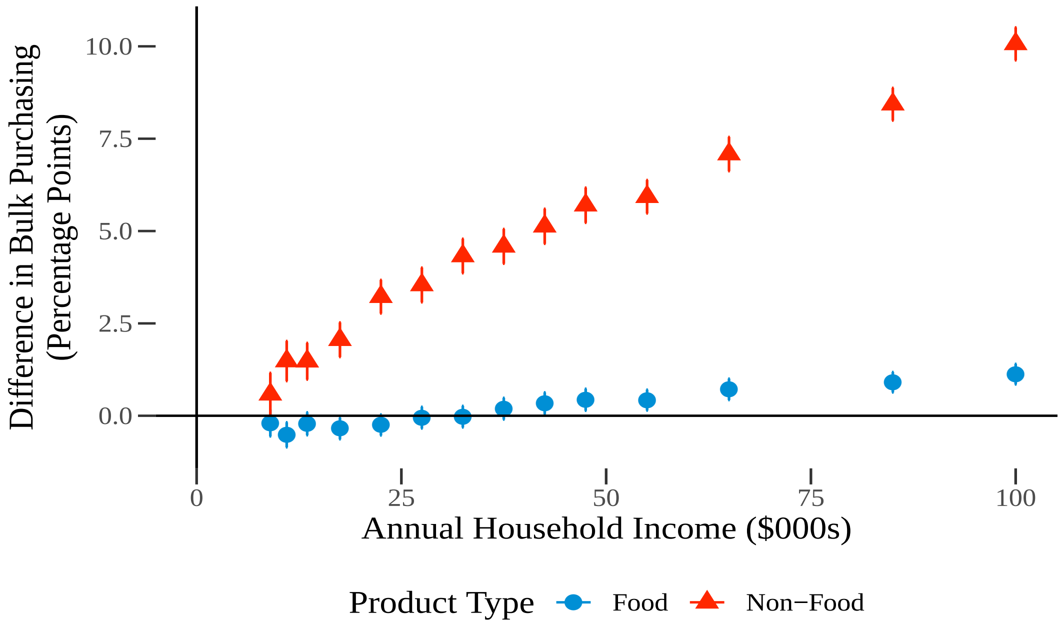
Note: Using Nielsen Retail Scanner data from 2016, this figure plots the distribution of coefficients from a regression of log unit price on log package size (Equation (1.1)) for individual product categories. Regression controls for store-brand-week fixed effects. Histogram plots 645 product categories.

food and non-food items. Because of these differences, I analyze food and non-food products separately. Following the literature, I classify a product as "bulk" if it is in the top two quintiles of the size distribution for that product category (Griffith et al. 2009).[8] Then, for each household, I compute the expenditure share of bulk purchases of food and non-food items. I then regress this "bulk share" on household income and other household characteristics that could affect consumption patterns and may be correlated with income, and plot the income coefficients. The equation below is estimated on food and non-food purchases separately:

$$BulkShare_{imt} = \alpha + \sum_{q} \beta^q Income_{imt} + \gamma X_{imt} + \lambda_m + \lambda_t + \epsilon_{imt}, \qquad (1.2)$$

[8] This definition avoids the risk of too narrowly defining bulk and only capturing purchases that occur solely at warehouse clubs. This broader definition helps capture large sizes that are available at grocery stores and mass merchandisers.

Figure 1.2: Bulk Purchasing by Household Income and Product Type



Note: Using 2004–2017 Nielsen Consumer Panel data, this figure plots the income bin coefficients from Equation (1.2), which regresses the share of annual purchases that were bulk packages on household characteristics as well as market and year fixed effects. Nielsen projection weights are used to ensure national representativeness. Households making $5–8k are the reference group.

where $BulkShare_{imt}$ is household i's share of bulk purchases in market m in year t (a market is a county). $Income_{imt}$ consists of dummies for each income bin q. X_{imt} consists of household characteristics (age, household composition, marital status, education, housing type, tract-level vehicle access).[9] Year and market fixed effects are captured by λ_t and λ_m.

Figure 1.2 illustrates that bulk purchases compose a 10 percentage point larger share of non-food expenditures for households making over $100,000 compared to those making $5,000–$8,000. As income increases, bulk purchases make up an increasing share of expenditures. For food items, there is a more muted increase of one percentage point across income groups.

The 10 percentage point gap is quite large. For the average household making

[9]These characteristics are used consistently throughout the paper. See Appendix 1.A for details of demographic variables and how they are collected.

between \$5,000-8,000, 39.6% of their non-food grocery spending is on bulk packages. Hence, households making over \$100,000 are 26% more likely to buy in bulk relative to the lowest-income group.

These patterns are consistent with high-income households buying in bulk, obtaining low unit prices, and consuming out of storage. Given the existence of quantity discounts, larger packages generally correspond to lower unit prices. The fact that low-income households are less likely to buy these storable items in bulk suggests that some obstacles may prevent them from buying and storing large packages.[10]

Because the bulk buying gap is largest for non-food products, the rest of this paper focuses on non-food products. These products are ideal for analyzing bulk purchasing because they isolate the key features that make bulk buying and quantity discounts attractive for households. Primarily, households can store items for future consumption. Additionally, these products generally do not have substitutes and they cannot be produced at home (e.g., toilet paper, diapers, etc.). My findings carry over to food products, but one must be careful to account for perishability, which counteracts product storability. Additionally, many food products have close substitutes (e.g., soda, juice, water, etc.) and home production (e.g., cooking meals) can substitute for many products (Aguiar and Hurst 2005; Aguiar and Hurst 2007).

Savings from Bulk Buying

In this subsection, I calculate the savings that low-income households could achieve from buying in bulk like high-income households. For each product category, I compute the average difference in package sizes by estimating the following regression:

$$\ln(AvgSize)_{imt} = \alpha + \sum_q \beta^q Income_{imt} + \gamma X_{imt} + \lambda_m + \lambda_t + \epsilon_{imt}, \tag{1.3}$$

where $AvgSize_{imt}$ is the quantity-weighted average package size purchased by household i in market m in year t, where a market is a DMA.[11] $Income_{imt}$ is an indicator for a household's income quartile. X controls for household characteristics. Market and year fixed effects are included through λ_m and λ_t.

[10] This relationship persists across most categories. Appendix 1.A shows the same pattern for a few popular categories and Figure 1.8 illustrates the difference for all non-food categories.

[11] Average package size is weighted by quantity to account for the fact that an unweighted average would favor small packages.

In this regression, β^q gives the average log-difference between the package size purchased by a household in income quartile q and the lowest-income quartile (households making less than \$25,000).[12] To compute savings, I multiply this average difference in package size purchased by the category-specific quantity discount estimated in Section 1.3. For example, high-income households buy 30% larger packages of toilet paper which has a quantity discount of 0.216. Therefore, low-income households could save $0.3 \times 0.216 = 0.0648$ or 6.5% from buying big packages like high-income households do. Aggregating across all categories where high-income households buy larger packages gives an estimated savings of 5%, or \$215, per year.[13,14]

Saving 5% on these common household purchases is substantial for low-income households. For the bottom quintile of the income distribution, these items account for 30% of their discretionary spending compared to 19% for the top quintile of the distribution.[15] If the about 24.4 million households making under \$25,000 were to obtain these savings, that would be an overall savings of \$5.4 billion annually, assuming no supply-side changes.[16] For context, this is equal to 8% of the \$68 billion federal Supplemental Nutrition Assistance Program budget in 2017 (USDA 2019). These potential savings do not require low-income households to buy more over the course of the year because buying in bulk does not necessarily change how much households *consume*. It just changes how much they *buy* at one time.

[12] I use quartiles to reduce the number of income bins from 15 to 4, but results hold at more granular levels. Disaggregated results are available upon request.

[13] This averages only across categories where high-income households buy larger packages. There are some categories, such as septic tank cleaners, in which high-income households buy in smaller packages. Imposing that low-income households buy the same average size across *all* categories reduces projected savings to 2.3%.

[14] The first-best calculations of savings would identify the product with the lowest unit price given a household's brand and store choice and compute savings based on that product. This estimate will likely be substantially higher than what I computed, so I view the estimated 5% savings as a conservative estimate of potential savings. See Appendix 1.A for calculations of savings on popular product categories.

[15] Discretionary spending is defined as total expenditures minus expenditures on shelter, utilities, transportation, healthcare, cash contributions, personal insurance, and pensions. Calculation is based on expenditure data on food at home and housekeeping supplies from Table 1 of the 2017 Consumer Expenditure Survey available at https://www.bls.gov/opub/reports/consumer-expenditures/2017/home.htm

[16] Household count comes from Table B19001 of the 2017 1-year American Community Survey.

1.4 Factors Affecting Bulk Buying

In this section, I show that *cognitive costs*, *store preferences*, *budget constraints*, and *storage costs* affect the bulk buying decision. To do this, I use plausibly exogenous variation and natural experiments to estimate the causal impact of unit pricing regulation and warehouse club entry on bulk purchasing. Since the biggest differences in bulk buying are for non-food grocery items, all analysis is restricted to non-food products.

Cognitive Costs

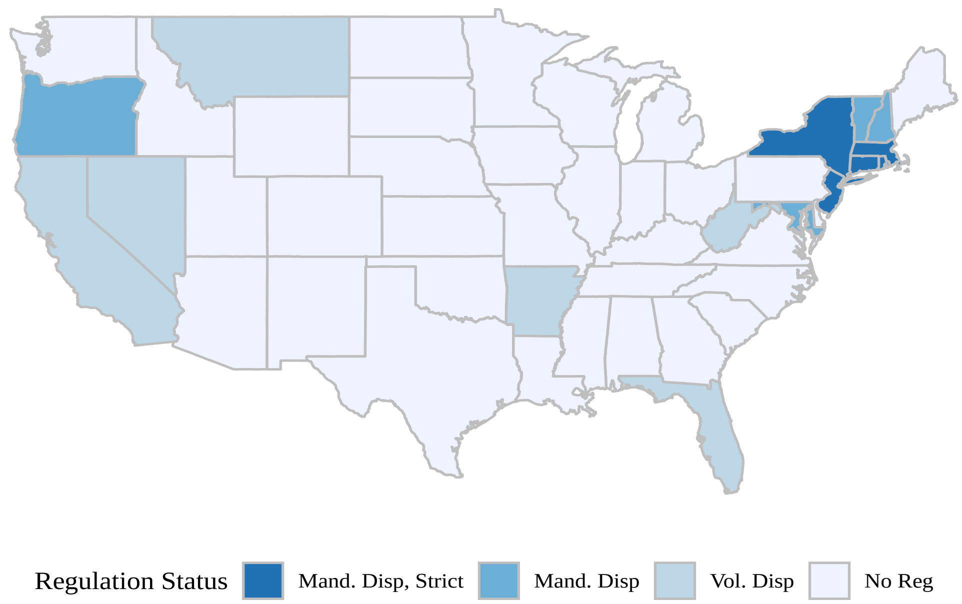
Cognitive costs are the first possible contributor to the bulk buying decision. Consumers may not be aware of the quantity discount (or how valuable it is) because they do not compute unit prices when making purchases. To test this hypothesis, I utilize a novel hand-collected dataset of state-level unit-price regulations requiring retailers to display per-unit prices. Displaying per-unit prices reduces cognitive costs and households can more easily compare products and pick the one with the best value.

Unit price labeling dates back to the late 1960s and early 1970s. During this period, a large consumer protection movement pushed for unit prices to be posted so consumers could compare different brands and sizes of products (Miyazaki, Sprott, and Manning 2000). As a result, some states passed laws requiring retailers to display unit prices. These laws varied widely with some giving retailers discretion over how to display unit prices and other states specifying formatting requirements, such as minimum font sizes and background colors to aid readability and clarity (Rose 2000).

Using annual regulatory updates published by the National Institute of Standards and Technology (NIST), I compile state-level regulations on unit pricing (NIST 2019). For states with regulations, I cross-check NIST's designation with state regulatory codes and consult with state officials to ensure accuracy. Figure 1.3 shows that, as of 2017, 16 states have regulations on the display of unit prices and 34 have no regulations.[17]

[17] Summary statistics of these groups are reported in Appendix Table 1.22.

Figure 1.3: Unit Price Regulations by State (2017)



Note: Using data from NIST Handbook 130, this figure plots whether or not a state has regulations in place governing the display of unit prices as of August 1, 2017. “No Reg” denotes that no regulations are in effect. “Vol. Disp” denotes states where regulations apply if retailers choose to display unit prices. “Mand. Disp” denotes states where all retailers must display unit prices. “Mand. Disp, Strict” denotes states where strict display formatting requirements are in effect.

If these regulations affect household decisions, then bulk buying should differ between states with and without these regulations. I first document how aggregate patterns in bulk buying differ between states with different regulations and then I will provide causal evidence for the impact of these regulations. I estimate the following regression:

$$BulkShare_{it} = \alpha + \beta_1 Reg_{it} + \gamma X_{it} + \lambda_t + \epsilon_{it}, \tag{1.4}$$

where $BulkShare_{it}$ is the annual share of expenditures that were bulk purchases for household i in year t. Reg_{it} is an indicator for whether or not unit-price regulations are in effect. X_{it} controls for household characteristics. I control for year fixed effects through λ_t. Standard errors are clustered by state because these regulations are at the state level.

Since 2004, no state has modified its regulations on unit prices, so the coefficient on unit pricing regulation is identified from cross-sectional variation between states that have regulations and those that do not.[18,19] Columns (1) and (2) of Table 1.2 reveal that bulk purchasing is 3.6 percentage points higher in states with unit price regulations compared to states without unit price regulation, even after controlling for household characteristics and year fixed effects.

Table 1.2: Unit Price Regulations and Bulk Buying

	(1)	(2)	(3)	(4)
Regulation	0.035**	0.036**		
	(0.017)	(0.016)		
Vol. Disp			0.050*	0.011
			(0.027)	(0.010)
Mand. Disp			0.038***	0.038***
			(0.009)	(0.009)
Mand. Disp, Strict			0.028***	0.028***
			(0.006)	(0.006)
Avg Bulk	0.5	0.5	0.5	0.49
Demographics	N	Y	Y	Y
Omit California	N	N	N	Y
Observations	732,512	732,512	732,512	668,791
Adjusted R^2	0.006	0.057	0.045	0.037

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Using Nielsen 2004–2017 Consumer Panel data combined with state-level regulations, this table shows the results of estimating Equation 1.4. The dependent variable is the annual share of bulk purchases made by households and the independent variables are either a binary indicator of the presence or absence of regulation (Columns (1) and (2)) or an ordered measure of regulatory stringency (Columns (3) and (4)). Column (4) omits California because it is the only state that has voluntary unit price, but strict requirements on how unit prices are displayed. Standard errors are clustered by state.

I then analyze these unit pricing regulations at a higher level of detail. State regulations vary across two dimensions: Posting and Formatting. Table 1.3 shows the breakdown of states along these dimensions. First, states can opt to have unit

[18] Because there is no time variation in regulations, I cannot include state fixed effects in the estimation.

[19] In 2013, the District of Columbia passed a law requiring retailers to display unit prices, but no households in my sample live in DC.

price posting be voluntary (seven states) or mandatory (nine states). Second, states can specify how unit prices are formatted when they are displayed.[20] Formatting regulations specify features including minimum font sizes, background colors, and label positioning. With the exception of California, only states that mandate unit price posting have formatting requirements. Excluding California, regulations are naturally ordered: no regulation, voluntary posting, mandatory posting (no formatting requirements), and mandatory posting (with formatting requirements).

Table 1.3: Unit Price Regulations by State

	No Formatting Rules		Strict Formatting Rules	
Voluntary Posting	Arkansas	Montana	California	
	Florida	Nevada		
	Hawai'i	West Virginia		
Mandatory Posting	Maryland	Vermont	Connecticut	New York
	New Hampshire		Massachusetts	Rhode Island
	Oregon		New Jersey	

Note: Based on state regulatory codes, the above table reports whether unit price posting is mandatory or voluntary for retailers and whether or not there are strict formatting requirements on how unit prices should be displayed (minimum font size, color, etc.).

Columns (3) and (4) continue the earlier analysis, but leverage the stringency of the regulations. Column (3) shows that mandatory posting is associated with significantly higher bulk buying, but states with voluntary requirements may have higher rates of bulk buying. However, as Table 1.3 shows, California is an outlier in this regulatory environment because is the only state with the unique combination of voluntary posting and strict formatting requirements. Because of this, I exclude California and re-estimate the regression. Column (4) reveals that California is the primary driver of this effect and states with voluntary posting do not have significantly higher bulk purchasing. On the other hand, mandatory unit price posting is associated with a 2.8–3.8 percentage point increase in bulk buying. The point estimates for bulk buying in states with strict formatting requirements are lower than those in states without formatting requirements, but these estimates are not significantly different from each other. This pattern supports the intuition that standardized unit price

[20]All states with these regulations standardize how unit prices are to be calculated, which is what makes the voluntary states different from states without regulations.

presentation reduces cognitive costs, increases the salience of unit prices, and facilitates comparisons for consumers.

This estimation provides strong evidence of a relationship between unit pricing regulations and bulk purchasing. However, there is a risk of selection bias since these regulations were primarily adopted in the Northeast and West Coast regions of the United States. To provide causal evidence, I examine about 13,000 households that move once during their tenure in the data. About 11% of these households move between regulatory regimes while the remainder are either local moves or moves that maintain their current regulatory regime. To estimate the effect of unit-price regulations on these two groups of movers, I use a differences-in-differences specification:

$$BulkShare_{it} = \alpha + \beta_1 Reg_{it} + \gamma X_{it} + \lambda_i + \lambda_t + \epsilon_{it}, \tag{1.5}$$

where the variables are the same as in Equation (1.4), but I control for household fixed effects and standard errors are clustered at the household level.[21] With this specification, β_1 is identified by changes in bulk purchases for households that move from a state with unit-price regulations to a state without unit-price regulations (or vice versa).[22] Since the "direction" of a household's move may matter (i.e., whether they start in a state without regulations and move to a state with regulations or vice versa), in subsequent specifications, I will account for the direction of the move.[23]

This specification relies on the assumption that households would have continued buying in bulk like other households that moved, but did not change their regulatory regime. To provide evidence supporting this "common trends" assumption, I plot an event study by estimating a modified version of Equation (1.5):

$$BulkShare_{it} = \alpha + \sum_{\tau \neq -1} \beta_1^{\tau} Yr_{it} + \gamma X_{it} + \lambda_i + \lambda_t + \epsilon_{it}, \tag{1.6}$$

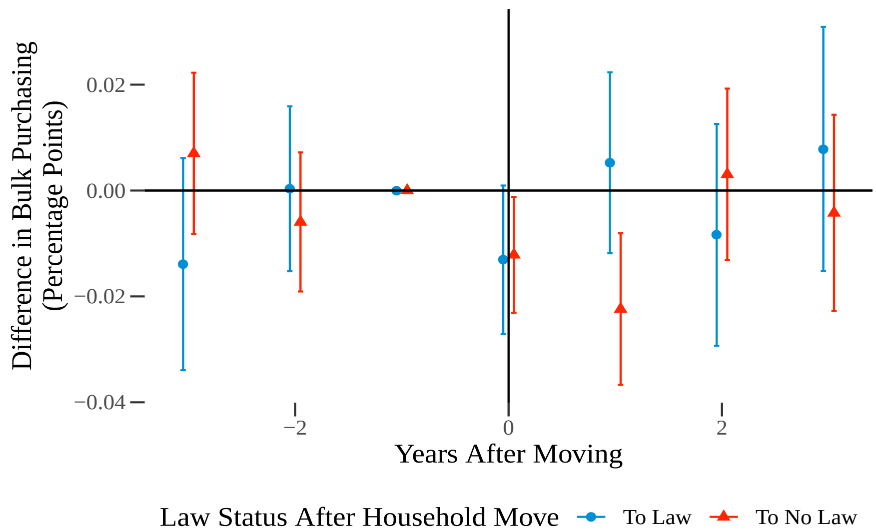
where Yr is a dummy for each year before or after a household moves to a state with a different unit pricing regime. The reference group is $t = -1$ so all effects are relative to the year before the household moves. Figure 1.4 plots the annual coefficients.

[21] Clustering at the state level does not affect the estimates.

[22] Projection weights are not used because the weights are not designed for this subsample of movers.

[23] Appendix Table 1.23 reports summary statistics for various groups of movers. Groups are relatively similar, but movers are slightly richer, older, more educated, and have fewer children.

Figure 1.4: Event Study of Movers and Unit Price Regulations



Note: Using 2004–2017 Nielsen Consumer Panel data, this figure plots the β_1^t coefficients and 95% confidence intervals from Equation (1.6), which regresses household bulk purchasing on dummies for years before and after a household moves to a state with a different unit pricing regime than the state it moves from. The regression controls for household characteristics as well as household and year fixed effects. Standard errors are clustered at the household level. "To" reports estimates for households that move from a state without unit price regulations to a state with unit price regulations. "Away" reports estimates for households that move from a state with unit price regulations to a state without regulations.

Figure 1.4 shows that there are no significant pre-trends. Furthermore, households decrease their bulk buying when they move from a state with unit-price regulations to a state without unit-price regulations. On the other hand, households that move from states without unit-price regulations to states with unit pricing regulations do not significantly change their bulk buying.

Table 1.4 reports the results of estimating Equation (1.5). Columns (1) and (2) show that a household's bulk buying is about one percentage point higher when they are in a state with unit price regulations, but this effect is only marginally significant. This specification implicitly assumes that the effect of moving to a state with unit price regulations will be the same as moving to a state without regulations (i.e., the effect is symmetric). Column (3) treats the different directions of moving differently and shows that moving to a state without unit price regulations significantly decreases bulk buying by 1.4 percentage points while moving to a state with regulations does not significantly change bulk buying.

Table 1.4: Event Study of Movers to Different State Regulatory Regimes

	All Movers		No Law To Law	Law To No Law
	(1)	(2)	(3)	(4)
Regulation	0.008	0.009*	0.003	0.014**
	(0.005)	(0.005)	(0.007)	(0.006)
Avg Bulk	0.5	0.5	0.5	0.5
Household FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Demographics	N	Y	Y	Y
Observations	92,739	92,739	86,712	88,479
Adjusted R^2	0.625	0.627	0.627	0.628
Note:	*p<0.1; **p<0.05; ***p<0.01			

Note: Using 2004–2017 Nielsen Consumer Panel data and state-level regulations, this table shows estimates of Equation 1.5 which regresses household bulk buying on unit price regulation after controlling for household fixed effects. "Regulation" denotes the estimated effect of moving from a state without regulation to a state with regulation. Columns (3) and (4) only include one set of movers that switch regimes in each specification with the remaining households that move, but do not switch regimes. Column (3) excludes households that move to states without unit price regulations to restrict identification of the regulatory effect to households moving to states with regulations. Column (4) excludes households that move to states with unit price regulations to restrict identification of the regulatory effect to households moving to states without regulations. Standard errors are clustered at the household level.

The asymmetric effect of unit pricing indicates the importance of both cognitive costs and consumer education. For households that move to states without regulation, the negative coefficient suggests that cognitive costs are discouraging households from buying in bulk. For households that move to states with regulation, they may not know how to best use the information provided and therefore consumer education may help them recognize the value of quantity discounts and buy in bulk more.

Unit pricing regulations are relatively simple to implement for both policymakers and retailers. Retailers will bear some initial setup costs of redesigning their price labels, but ongoing costs will likely be similar to current menu costs that firms bear.[24]

[24] In 1975, the Government Accountability Office (then the General Accounting Office) estimated

Adopting unit pricing policies (like those recommended by the National Conference on Weights and Measures) would encourage bulk buying while imposing few costs. These findings support the broader assertion that increasing price transparency allows households to choose products that deliver more value.

Store Preferences

The second potential contributor to the bulk buying gap is *store preferences*; low-income households may not live in areas where bulk sizes are available or may not shop at stores that offer bulk sizes. In this subsection, I provide evidence that the bulk buying gap persists within neighborhoods and within store types. Then, I show that warehouse club entry increases bulk buying by 4.0–7.3%, but these increases hold only for middle- and high-income households.

Inequality Within Markets and Retail Chains

If supply factors are the primary driver of the bulk buying gap, then the gap should disappear when comparing households in the same neighborhood since they have the same set of stores to choose from. I show that the bulk buying gap still persists within ZIP codes. This remaining gap corresponds to the amount that *cannot* be explained by differences in access, at least as approximated by geography.

Even within ZIP codes, there may be other factors affecting where households shop, such as whether or not a household has a vehicle, access to public transit, or a warehouse club membership. To account for possible differences, I examine how much of the bulk buying gap persists within chains. This exercise assumes that within a chain, households have access to the same assortment of goods (DellaVigna and Gentzkow 2019). I also examine the bulk buying gap within store types (i.e., "channel") to account for the fact that bulk buying differences may primarily be between channels (discount versus dollar) instead of between retailers within a channel (Walmart versus

that implementation and maintenance would cost about 0.1% of sales (General Accounting Office 1975). This was estimated before the adoption of bar codes and other efficiency-improving practices of the retail sector. Implementing unit pricing now is likely to cost substantially less than those early estimates.

Target).[25]

I estimate within-ZIP and within-chain bulk buying gaps using a modified form of Equation 1.2:

$$BulkShare_{imt} = \alpha + \sum_q \beta^q Income_{imt} + \gamma X_{imt} + \lambda_{mt} + \epsilon_{imt}, \tag{1.7}$$

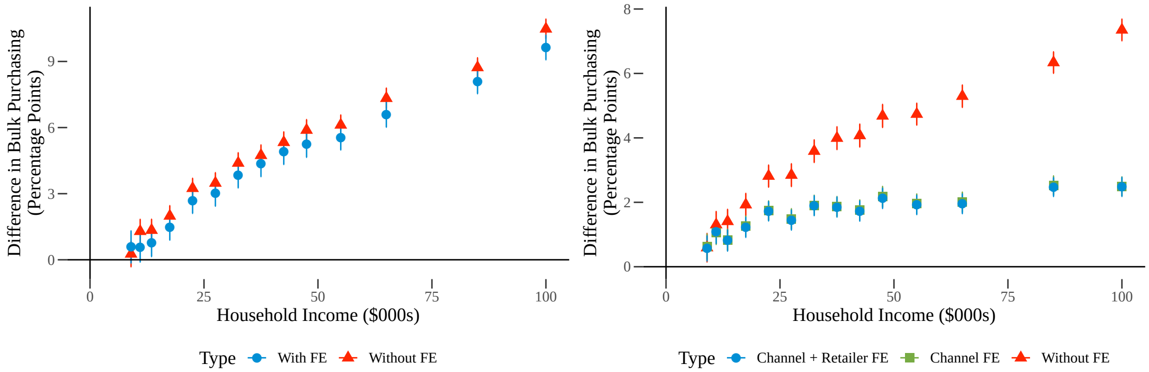
where $Income_{imt}$ is an indicator for the income bin of household i in market m in year t. X_{imt} consists of household characteristics. For the analysis of bulk buying within ZIP code, $BulkShare_{imt}$ is the share of bulk purchases made by household i in ZIP code m in year t and λ_{mt} is a ZIP-year fixed effect. For the analysis of bulk buying within retail chains, $BulkShare_{imt}$ is the share of bulk purchases made by household i in retail chain m in year t and λ_{mt} is a retail chain-year fixed effect and/or a channel-year fixed effect.

Figures 1.5a and 1.5b plot the income coefficients with and without fixed effects for each regression. Adding ZIP-year fixed effects reduces the gap between the highest and lowest income groups by 9% (from 10.5 percentage points to 9.6 percentage points). Results are virtually unchanged if I use county-year fixed effects instead of ZIP-year fixed effects. Using channel-year fixed effects reduces the bulk buying gap by a more substantial 66% (from 7.4 percentage points to 2.5 percentage points). Adding retail chain-year fixed effects on top of channel-year fixed effects does not significantly affect the bulk buying gap. This implies that a large share of the bulk buying gap is related to the *types* of stores households shop at, but not the specific chain they choose within a particular store type.

Overall, within ZIP codes, the bulk buying gap between high- and low-income households persists. However, within store type (or retail chain), the bulk buying gap is substantially reduced. Two important conclusions can be drawn from these patterns. First, in an accounting sense, the type of store a household shops at accounts for two-thirds of the bulk buying gap. This is likely an overestimate of the contribution of store preferences because those preferences may be driven by more fundamental factors (e.g., high storage costs or budget constraints could prevent households from shopping at warehouse clubs, as opposed to households having a low preference for

[25] Retailer names are only for expository purposes. Retailer identities are anonymized in the Nielsen data.

Figure 1.5: Bulk Buying Within ZIP Codes and Within Retailers



(a) ZIP Code FEs

(b) Retailer / Channel FEs

Note: Using 2004–2017 Nielsen Consumer Panel data, panel (a) plots the income bin coefficients from Equation (1.7), which regresses the share of annual purchases that were bulk packages on household characteristics as well as a ZIP code-year fixed effect. Panel (b) runs the same estimation, but with retailer-year and channel-year fixed effects (a "channel" is a type of store). Nielsen projection weights are used to ensure national representativeness. Households making $5–8k are the reference group.

warehouse clubs). Second, the bulk buying gap *still persists* within channels and retail chains. These patterns suggest that where a household shops and what they choose within a store are much more important than where a household is located. The next section explores how store preferences are related to income and how warehouse clubs affect bulk buying.

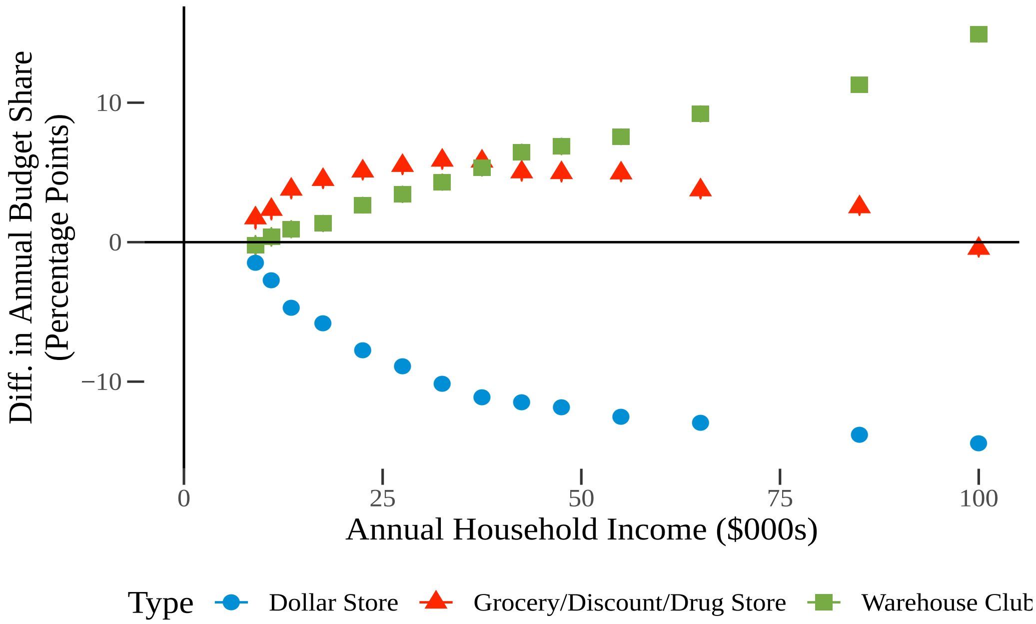
Store Preferences by Income

The previous section shows that while the bulk buying gap persists within ZIP codes, it is narrower within store types and retail chains. In this section, I show that the biggest shopping differences between income groups are related to warehouse clubs. I then estimate the effect of warehouse club entry on bulk buying.

To demonstrate differences in store preference by household income, I examine the relationship between where households shop and their income using the following regression:

$$ChannelShare_{imt} = \alpha + \sum_q \beta^q Income_{imt} + \gamma X_{imt} + \lambda_m + \lambda_t + \epsilon_{imt}, \tag{1.8}$$

Figure 1.6: Annual Spending Shares By Store Type



Note: Using 2004–2017 Nielsen Consumer Panel data, this figure plots the income bin coefficients from Equation (1.8), which regresses the share of annual purchases at each store type on household characteristics as well as year and market fixed effects. Nielsen projection weights are used to ensure national representativeness. Households making $5–8k are the reference group.

where $ChannelShare_{imt}$ is the share of annual spending that household i in market m in year t made in a particular channel (grocery store, discount store, dollar store, drug store, or warehouse club). $Income_{imt}$ is an indicator for a household's income bin. X_{imt} captures other household characteristics. Finally, market and year fixed effects capture differences in spending shares across markets and over time.

Figure 1.6 reveals that while there are small differences in the share of annual expenditures at grocery, drug, and discount stores, there are dramatic differences in whether households shop at warehouse clubs or dollar stores: households making over $100,000 spend about 13 percentage points more of their non-food expenditures at warehouse clubs than households making under $25,000.

Because the biggest differences are in warehouse clubs and these stores almost solely stock bulk sizes, I focus on how warehouse clubs affect bulk buying. The following analysis of warehouse clubs uses data on over 1,400 warehouse club locations between 2004–2015.[26] The first possibility is that high-income households shop at warehouse clubs because they are closer. Table 1.5 shows that low-income households are about 15 miles away from the nearest warehouse club compared to only 8 miles away for high-income households.

Table 1.5: Average Distance to Warehouse Club by Income (Miles)

Household Income	Mean	SD	25th Pctile	75th Pctile
<25k	14.79	18.98	3.26	20.53
25-50k	12.79	17.51	3.03	15.94
50-100k	10.61	15.16	2.85	12.01
>100k	7.92	12.08	2.51	8.45

Note: Using Nielsen Consumer Panel data from 2004–2015 and warehouse location data, this table reports the distance between ZIP code centroids of warehouse locations and household locations. Nielsen projection weights are used to ensure national representativeness.

The ideal experiment would randomly assign warehouse clubs to neighborhoods and then their effect on bulk buying could easily be calculated. Even though store locations are not randomly assigned, within a household, it is exceedingly unlikely that a warehouse club opening could be co-incident with a shift in bulk buying, so any observed changes are likely causal. Leveraging the panel structure of the Nielsen Consumer Panel, I estimate how a household's bulk purchasing changes after a warehouse club opens using the following equation:

$$BulkShare_{imt} = \alpha + \beta Entry_{imt} + \gamma X_{imt} + \lambda_{im} + \lambda_t + \epsilon_{imt}, \tag{1.9}$$

where $BulkShare_{imt}$ is the share of bulk purchases made by household i in market m in quarter t. $Entry_{imt}$ is an indicator for whether or not a warehouse club entered within 15 miles of household i in quarter t.[27,28] I include a household-market and year-quarter fixed effects λ to ensure that β is identified by within-household changes

[26] Data provided by the authors of Coibion, Gorodnichenko, and Koustas (2017) and covers BJ's, Costco, and Sam's Club.

[27] In cases where a household is located near multiple warehouse clubs, I use the earliest entry date since the first warehouse club would generate the largest supply shock.

[28] According to the 2017 National Household Travel Survey, the average household traveled about seven miles to buy goods, with low-income households traveling about one or two miles less than higher-income households (Federal Highway Administration 2017). Allowing for the possibility that households might travel farther to shop at a warehouse club, I use a cutoff of 15 miles. Appendix Table 1.24 shows that this pattern is robust to other cutoffs.

in bulk buying before and after a warehouse club opens instead of households that may move to areas closer to warehouse clubs. X controls for possible demographic changes within the household.

This specification relies on the assumption that households would have continued buying in bulk like other households that did not experience a warehouse club entry. To provide evidence supporting this "common trends" assumption, I plot an event study by estimating a modified version of Equation 1.9, but replace the entry indicator with dummies for each quarter pre- and post-entry:

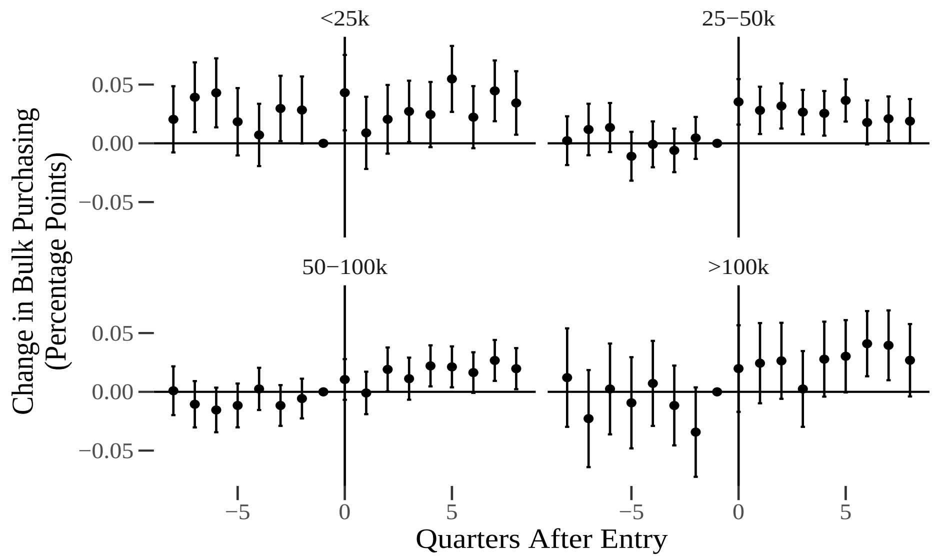
$$BulkShare_{imt} = \alpha + \sum_{q} \beta^{q} Qtr_{imt} + \gamma X_{imt} + \lambda_{im} + \lambda_{t} + \epsilon_{imt}, \tag{1.10}$$

where Qtr_{imt} is a dummy for each quarter prior to entry and after entry, with the quarter immediately before entry ($q = -1$) as the reference group. Figure 1.7 plots the quarterly coefficients and shows that for most income groups there are no significant pre-trends. For households in the lowest income quartile, there is some evidence that those that experienced a warehouse club entry buy in bulk more often than other low-income households that do not experience an entry. After a warehouse club enters, there are significant increases in bulk buying for middle- and high-income households and these effects are persistent up to eight quarters after a warehouse club has opened.

Table 1.6 shows the regression results. Overall, households that experienced a warehouse club entry increased their bulk purchasing by two percentage points. However, when I interact household income with warehouse club entry, the increase in bulk buying is due to changes for households making over $25,000 and is increasing in income, with households making over $100,000 increasing their bulk buying by 3.5 percentage points. Households in the lowest quartile do not have any significant change in their bulk buying. One likely reason that low-income households do not change their bulk buying is that even after a warehouse club enters, households do not purchase a membership (fees range from $45–$120 depending on the chain and membership level). Other possible reasons are that low-income households do not have access to transportation that can carry items home, do not have the space to store the items, or even if they had a membership, they still would not purchase extremely large sizes available at warehouse clubs due to budget constraints.[29]

[29] As an example, Philadelphia provides public transit access to a warehouse club. However,

Figure 1.7: Event Study of Warehouse Club Entry



Note: This figure plots the quarterly coefficients from Equation (1.10)—the effects of warehouse club entry on bulk purchasing of households before and after warehouse club entry—using 2004–2015 household-by-quarter Nielsen Consumer Panel data. The regression controls for household characteristics as well as household-ZIP code fixed effects. All coefficients are relative to bulk purchasing in the quarter before entry ($q = -1$). Error bars denote 95% confidence intervals.

This analysis estimates the intent to treat effect since not all households shop at the entrant warehouse club after it opens. As a result, this is a conservative lower bound on the actual treatment effect on households that shop at warehouse clubs.[30] The effect is quite substantial even given how conservative it is.

To examine whether this change in bulk buying comes from households shopping more at warehouse clubs, I estimate Equation 1.9 on different margins of bulk buying. The increase in bulk buying after warehouse club entry could be coming from three possible margins. First, households could increase their bulk buying at non-warehouse club stores (intensive non-warehouse club margin). Second, households could increase their bulk buying at warehouse club stores (intensive warehouse club margin). Third,

carrying club-sized items on a bus is infeasible for more than two or three items. A personal vehicle would be necessary.

[30]Even though low-income households do not change their bulk buying, other research suggests that they may be worse off because existing retailers are more likely to increase prices for storable products as a competitive response (Bauner and Wang 2019).

Table 1.6: Effect of Warehouse Club Entry on Bulk Buying

	(1)	(2)	(3)
Post-Entry	0.020***	0.020***	
	(0.004)	(0.004)	
Post-Entry : <25k			0.001
			(0.008)
Post-Entry : 25-50k			0.019***
			(0.006)
Post-Entry : 50-100k			0.025***
			(0.005)
Post-Entry : >100k			0.035***
			(0.009)
Avg Bulk	0.48	0.48	0.48
Household-ZIP FE's	Y	Y	Y
Year-Quarter FE's	Y	Y	Y
Demographic Controls	N	Y	Y
Observations	2,401,038	2,401,038	2,401,038
Adjusted R^2	0.428	0.428	0.428
Note:	*p<0.1; **p<0.05; ***p<0.01		

Note: This table uses 2004–2015 Nielsen Consumer Panel data at the household-quarter level. Coefficients are reported for Equation (1.9) which regresses households' quarterly bulk purchase shares on an indicator for warehouse club entry, household characteristics as well as household-ZIP code and year-quarter fixed effects. Projection weights are not used.

households could increase their bulk buying by switching shopping to warehouse club stores (extensive margin). The observations used in estimation are household-quarter bulk shares at warehouse club and non-warehouse club stores.

Table 1.7 shows the regression results. Column (1) is estimated using quarterly bulk buying shares at non-warehouse club stores as the dependent variable. After warehouse club entry, there is no significant change in bulk buying at non-warehouse club stores for any income group. Furthermore, the standard errors are quite small, so if there were changes, they are minimal. Column (2) is estimated using quarterly bulk buying shares at warehouse club stores as the dependent variable, which includes many zeros because there are quarters where households do not shop at warehouse club stores. For households that never shop at warehouse club stores, they have a

Table 1.7: Effect of Warehouse Club Entry on Bulk Buying Along Different Margins

	Non-Club	Club	Club Intensive
	(1)	(2)	(3)
Post-Entry	0.0001	0.075^{***}	-0.005
	(0.004)	(0.007)	(0.004)
Post-Entry : 25-50k	-0.003	0.018^{***}	-0.003
	(0.004)	(0.006)	(0.004)
Post-Entry : 50-100k	-0.001	0.009^{***}	0.0005
	(0.002)	(0.003)	(0.003)
Post-Entry : >100k	-0.001	0.012^{***}	-0.001
	(0.003)	(0.004)	(0.003)
Avg Bulk	0.4	0.27	0.97
Household-ZIP FE's	Y	Y	Y
Year-Quarter FE's	Y	Y	Y
Demographic Controls	Y	Y	Y
Observations	2,401,038	2,401,038	661,240
Adjusted R^2	0.254	0.567	0.193
Note:	*p<0.1; **p<0.05; ***p<0.01		

Note: This table uses 2004–2015 Nielsen Consumer Panel data at the household-quarter level. Coefficients are reported for Equation (1.9) which regresses households' quarterly bulk purchase shares on an indicator for warehouse club entry, household characteristics as well as household-ZIP code and year-quarter fixed effects. Column (1) uses bulk buying shares only at non-warehouse club stores as the dependent variable. Column (2) uses bulk buying shares at warehouse club stores (inclusive of zero spending quarters) as the dependent variable. Column (3) uses bulk buying shares at warehouse club stores (excluding zero spending quarters) as the dependent variable. Projection weights are not used.

string of zeroes. After warehouse club entry, quarterly bulk buying at warehouse clubs increases by a statistically significant 7.5 percentage points and is higher for higher income groups. This increase could be generated by households shopping more often at warehouse clubs, and therefore increasing the number of quarters with non-zero bulk spending at warehouse clubs. It could also be generated by households increasing the share of bulk purchases they already make at warehouse clubs. Column (3) focuses on only quarters with positive bulk shares at warehouse clubs and shows that there are no significant changes in bulk buying, conditional on shopping at a warehouse club, after a warehouse club enters. Therefore, the significant increase in Column

(2) is generated by households switching to shopping more at warehouse clubs after a warehouse club enters. Overall, warehouse club entry increases bulk buying by encouraging households to make more shopping trips to warehouse clubs.

Budget Constraints

Budget constraints may be another contributing factor to the bulk buying gap. Importantly, the necessary budget constraints must bind over short time periods (e.g., months) because over the course of a year, households spend more than they would have for the same amount of goods if they had taken advantage of bulk discounts. Such short-horizon budget constraints are most binding for the lowest income groups, but are unlikely to bind for middle- and high-income groups.[31]

To test this explanation, I examine over-the-month changes in liquidity. Low-income households are more likely to have higher liquidity at the start of the month compared to the end of the month (Stephens Jr 2003; Orhun and Palazzolo 2019). Consistent with this fact, the spending of low-income households tends to decline over the course of the month while the spending of higher income groups is relatively flat (Orhun and Palazzolo 2019).

I use within-household variation in the timing of purchases to estimate a differences-in-differences model. This model tests the coincidence of changes to bulk buying with times of the month when households may be liquidity constrained. This approach is similar to Orhun and Palazzolo (2019). I use weekly bulk buying information to estimate the following regression:

$$BulkShare_{iw} = \alpha + \sum_{q} \sum_{w=2}^{4} \beta_1^{qw} \mathbb{1}\{week = w\} * Inc_{iw} + \beta_2 Inc_{iw} + \gamma X_{iw} + \lambda_i + \epsilon_{iw}, \quad (1.11)$$

where $BulkShare_{iw}$ is the share of bulk purchases made by household i in week w. $\mathbb{1}\{week = w\}$ is an indicator for the second, third, or last week of the month. Inc_{iw} indicates the household's income bin q. X_{iw} is the usual set of household characteristics and λ is a household fixed effect. The results are reported in Table 1.8.

Overall, households making less than $25,000 decrease their bulk buying by a slight 0.2 to 0.4 percentage points during the third and fourth weeks of the month while

[31] Middle- and high-income households may still face monthly budget constraints, but grocery spending is unlikely to be a major factor for all but the lowest-income households.

Table 1.8: Over-the-Month Changes in Bulk Buying

	(1)	(2)	(3)
Week 2	-0.002^{***}	-0.001	-0.001
	(0.0004)	(0.001)	(0.001)
Week 3	-0.002^{***}	-0.004^{***}	-0.004^{***}
	(0.0004)	(0.001)	(0.001)
Week 4	0.002^{***}	-0.002^{**}	-0.002^{**}
	(0.0004)	(0.001)	(0.001)
25-50k		0.003^{***}	0.002^{**}
		(0.001)	(0.001)
Week 2 : 25-50k		-0.002^{**}	-0.002^{**}
		(0.001)	(0.001)
Week 3 : 25-50k		0.001	0.001
		(0.001)	(0.001)
Week 4 : 25-50k		0.003^{***}	0.003^{***}
		(0.001)	(0.001)
50-100k		0.009^{***}	0.004^{***}
		(0.001)	(0.001)
Week 2 : 50-100k		-0.001	-0.001
		(0.001)	(0.001)
Week 3 : 50-100k		0.002^{*}	0.002^{*}
		(0.001)	(0.001)
Week 4 : 50-100k		0.005^{***}	0.005^{***}
		(0.001)	(0.001)
>100k		0.025^{***}	0.012^{***}
		(0.001)	(0.001)
Week 2 : >100k		0.001	0.001
		(0.001)	(0.001)
Week 3 : >100k		0.006^{***}	0.006^{***}
		(0.001)	(0.001)
Week 4 : >100k		0.007^{***}	0.007^{***}
		(0.001)	(0.001)
Mean Bulk	0.46	0.46	0.46
Household FE's	Y	Y	Y
Demographics	N	N	Y
Observations	2,854,905	2,854,905	2,854,905
Adjusted R^2	0.402	0.402	0.404
Note:	*p<0.1; **p<0.05; ***p<0.01		

Note: Using 2004–2017 Nielsen Consumer Panel data, this table displays the regression coefficients from estimating Equation 1.11 which regresses a household's weekly share of bulk purchases of non-food products on the week of the month, income, and other household characteristics and includes a households fixed effect.

households making over $25,000 have no decline (and possibly a slight increase) in their bulk buying during the third and fourth weeks of the month.

This analysis does not rule out the possibility of liquidity constraints contributing to differences in bulk buying, but it does rule out that local changes in liquidity significantly affect bulk buying. Liquidity shocks larger than intra-month paycheck variation may be necessary to increase bulk buying, but more work would be necessary to determine whether that is the case.[32]

Storage Costs

Storage costs are the fourth contributing factor that I examine. Intuitively, households that buy in bulk need a place to store large packages, which could be in a basement, pantry, or cabinets. Households without available storage space may want to save money through quantity discounts, but choose not to because they have limited storage space.

The ideal experiment would randomly assign households to various home sizes and then observe their bulk purchasing behavior to identify storage costs. However, exogenously changing a household's living situation is infeasible. The next best option is to test some intuitive implications of storage costs. First, while I cannot randomly assign households to different home sizes, there are many households that move while they are in the Nielsen panel. I observe whether households live in single-family homes or apartments, which generates variation in available storage space. According to the American Housing Survey, the median single-family home is about twice as large as the median apartment. Since at least 1999, new single-family homes have had a median size of 2,000–2,400 square feet while the median apartment is only 1,000–1,100 square feet and this holds true within Census regions as well. Therefore, households that move into single-family homes are likely to have more available storage space and this will increase their willingness to buy in bulk.

To test this hypothesis, I estimate how bulk buying changes when households

[32] Another approach could leverage information on when households get paid. Households that get paid weekly or bi-weekly may not have such large monthly fluctuations over the course of the month compared to households that get paid monthly.

change their housing size, by estimating Equation 1.12:

$$BulkShare_{it} = \alpha + \beta House_{it} + \gamma X_{it} + \lambda_i + \lambda_t + \epsilon_{it}, \tag{1.12}$$

where $House_{it}$ is a dummy for whether a household i lives in a single-family house in year t (apartments are the reference group). X_{it} controls for changes in other household characteristics. Household and year fixed effects, λ, ensure that β is identified off of within-household changes in housing. Standard errors are clustered at the household level. Table 1.9 shows that bulk buying is one percentage point higher when households are in single-family homes compared to when they are in apartments.

Table 1.9: Relationship Between Bulk Buying and Housing Changes

	(1)	(2)	(3)
Single-Family Home	0.012***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)
Avg Bulk	0.5	0.5	0.5
Market FE's	Y	Y	Y
Year FE's	Y	Y	Y
Demographics	N	Y	Y
Future Income	N	N	Y
Observations	731,762	731,762	566,535
Adjusted R^2	0.688	0.688	0.691
Note:	*p<0.1; **p<0.05; ***p<0.01		

Note: Using Nielsen 2004–2017 Consumer Panel data, this table shows the results of estimating Equation 1.12. The dependent variable is the annual share of bulk purchases made by households and the independent variables are housing and other household characteristics. Estimation includes household, market, and year fixed effects. "Future Income" denotes a household's income one year in the future. Standard errors are clustered by household.

Column (1) shows that bulk buying is 1.2 percentage points higher when a household lives in a single-family home. However, since housing changes can be due to other within-household shifts, such as marriage or having children, column (2) also controls for other within-household demographic changes. The increase in bulk buying is slightly reduced, but there is still an increase when households move into larger spaces. Finally, households may move into larger housing if they expect to earn more and this expectation of future income may also increase their bulk buying. Column (3) also includes a household's one-year-ahead income and there is no change to the bulk buying increase after a household moves into a single-family home. This result is not causal, but it supports the intuition that when households have more storage space, they are more able to buy in bulk.

Another implication of storage costs is that products with a smaller "footprint" (physical volume) have lower storage costs. Therefore, if storage costs influence bulk buying, there should be a smaller gap in bulk buying for smaller products (like plastic wrap) relative to large, cumbersome products (like paper towels and toilet paper). To test this implication, I estimate a modified form of Equation 1.3 that relates average package sizes with household income:

$$\ln(AvgSize)_{imt} = \alpha + \sum_q \beta^q Income_{imt} + \gamma X_{imt} + \lambda_m + \lambda_t + \epsilon_{imt},$$

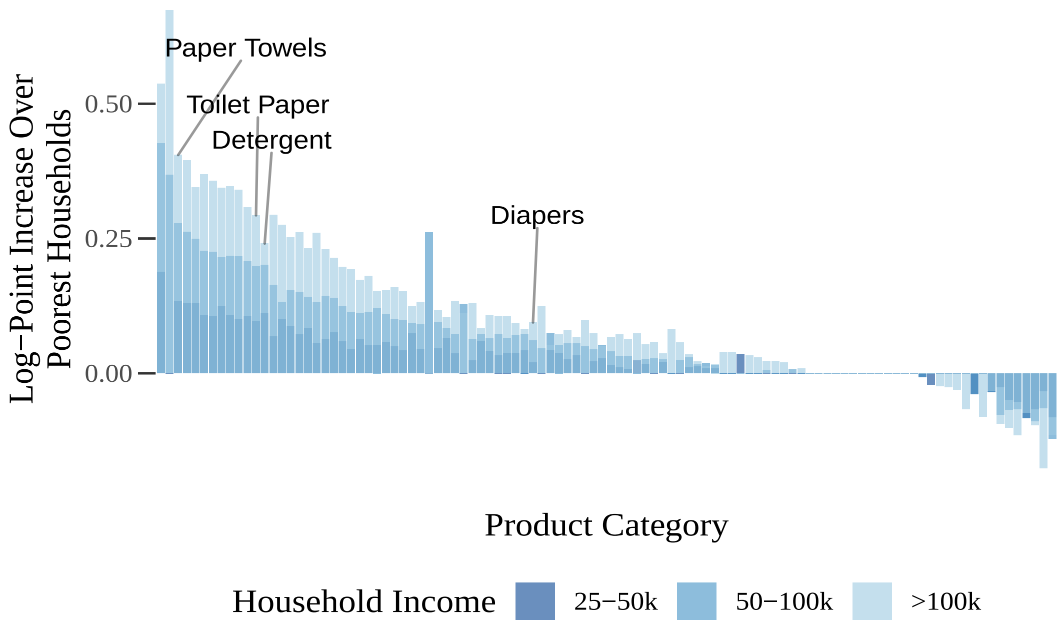
where $AvgSize_{imt}$ is the average package size purchased by household i in market m in year t. $Income_{imt}$ consists of dummies for each income quantile q. X_{imt} consists of household characteristics. Year and market fixed effects are captured by λ.

Figure 1.8 plots the income coefficients from the regression for all non-food grocery categories. I have highlighted some popular product categories. The bulk buying gap is largest for the physically biggest products such as paper towels and toilet paper while the gap is smaller for less bulky items such as liquid detergent and diapers. Overall, this pattern supports the hypothesis that storage costs contribute to the bulk buying gap, but the persistence of the gap even for smaller products suggests that other factors are at play. This graph should also be interpreted with caution because not all products are commonly consumed across all income groups. The two largest gaps are in pool cleaning supplies and paper bags, which are more likely to be purchased by high-income households than low-income households, regardless of package size.[33]

Overall, these results provide evidence that storage costs and bulk buying are related. When households move to larger homes (relative to apartments), they buy more in bulk. Similarly, product categories with larger physical footprints exhibit larger bulk buying gaps relative to product categories with smaller footprints. To more

[33] The gap in paper bag purchases may be related to preferences in that high-income households may use paper bags for lunches while lower-income households use a lunchbox or plastic bags.

Figure 1.8: Bulk Buying Gap For Non-Food Grocery Products



Note: Using 2004–2017 Nielsen Consumer Panel data, this figure plot the β coefficients from Equation (1.3), which regresses average package size purchased on household income. The regression controls for household characteristics as well as market and year fixed effects.

precisely quantify storage costs, I estimate a simple model of the consumer purchase decision.

1.5 Model

The previous analyses show that cognitive costs and storage costs affect the bulk buying decision. To decompose the contribution of each factor, I embed them into a discrete choice model of the household's purchase decision. The ideal setting would include a homogeneous good where demand is uncorrelated with income. Given substantial price, package size, and regulatory variation, differences in large and small purchases between households would identify storage costs and differences in buying between regulatory regimes would identify cognitive costs. This setting is approximated by one where products have limited dimensions of differentiation and storage costs can be separately identified from demand.

A discrete choice model of toilet paper purchases closely approximates this ideal setting. Toilet paper is an excellent product for this analysis because it is a necessity item with easily observable dimensions of differentiation, namely price, quality, quantity, and package size. It is offered in a wide range of package sizes and stores stock numerous brands and sizes (grocery and mass merchandise stores usually stock 35–40 unique brand-sizes). The top five brands and private-label store brands account for 86% of sales. I focus on the most common package sizes, which range from 4- to 24-roll packages. I define a product as a unique brand-size combination.[34] Additionally, underlying toilet paper consumption is primarily a function of household composition and age, not income.[35] High-income households consume a similar amount as low-income households but make fewer purchases (Orhun and Palazzolo 2019). Finally, toilet paper cannot be easily substituted for another product nor can it be obtained through home production.[36]

The biggest identification challenge is separately identifying storage costs from underlying demand (i.e., households may buy large quantities because they have high consumption or because they have low storage costs). To separate storage

[34]Specifically, this is a unique brand-roll count-sheet count because packages can differ in their "concentration" due to "double," "mega," and "super mega" rolls.

[35]A 100-fold cross-validated elastic net regression of annual purchases on household characteristics rules out income as significantly predictive. See Appendix 1.A for details.

[36]While a bidet is a possible alternative, this is more likely a lifestyle choice instead of a situation where households switch between toilet paper and bidets. Furthermore, in the United States, 98% of households report that they use toilet paper (the remainder either said no or did not respond) (Statista 2019).

costs from demand, I use variation induced by differences in product "concentration," which I define as the yield of the product per unit volume. Product concentration breaks the direct link between volume and consumption. In the detergent category, a product's yield is the number of washes it will supply. A concentrated detergent can wash the same number of loads but requires a smaller fluid volume than diluted detergent. Therefore, given the same number of washes, households that choose concentrated detergent must have higher storage costs than those choosing diluted detergent, assuming quality does not differ based on concentration.

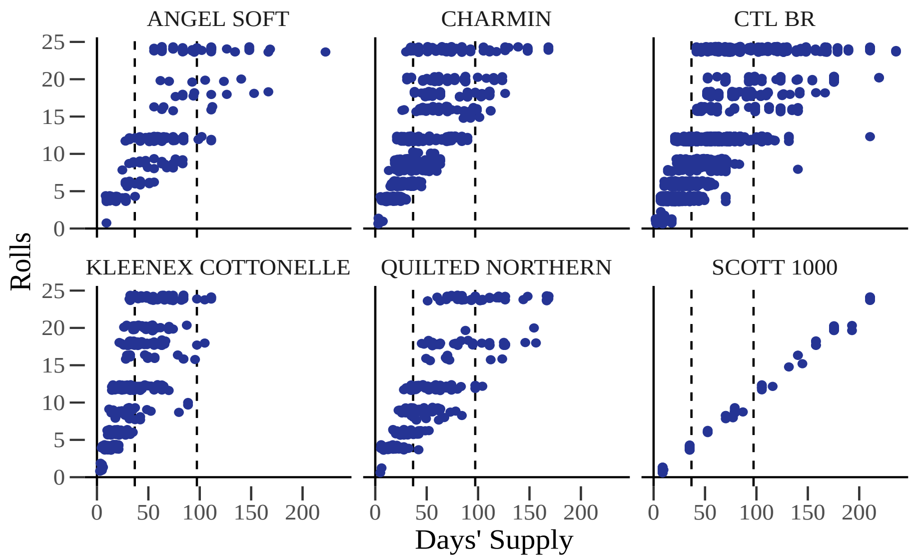
The same reasoning holds true for toilet paper. Households do not demand a particular number of rolls (the primary determinant of package size), but choose how long they want their supply to last (i.e., purchase enough to last for two weeks, a month, two months, etc.).[37] Toilet paper comes in a variety of concentrations with "mega" rolls being more concentrated than "regular" rolls. Therefore, a household that purchases 24 "regular" rolls has the same demand for toilet paper as a household that purchases six "mega" rolls, but the former household has lower storage costs since they can store the bigger package.

To illustrate the varying concentrations of toilet paper, Figure 1.9 plots the distribution of quantity (measured in number of days the supply will last for a single person) against package sizes (measured in rolls) for toilet paper products in the Nielsen data. As expected, there is an increasing relationship between how long the package will last and the number of rolls in a package, but there is substantial variation within packages containing the same number of rolls. The dashed lines denote the 25th and 75th percentiles of the average days' supply purchased by households. A wide range of package sizes fall within this range for each brand.[38] For example, a household demanding a 60-day supply of Charmin could purchase a package containing anywhere from 8 to 24 rolls. This overlap generates the necessary variation to separate storage costs from underlying demand.

[37]According to a 2007 Charmin survey, the average person uses 57 sheets per day. I assume this consumption rate when computing how long a product will last (Jaffe 2007).

[38]Scott toilet paper is an exception because it does not offer different roll types. All rolls have 1000 sheets.

Figure 1.9: Scatterplot of Toilet Paper Package Size and Quantity



Note: Using 2004–2017 Nielsen Consumer Panel data, this figure plots the package sizes and quantities of the top five toilet paper brands and private-label products. The y-axis represents the number of toilet paper rolls contained in a package while the x-axis represents the number of days a product will last a single person household assuming a consumption rate of 57 two-ply sheets per day (Jaffe 2007). Noise is added vertically to better illustrate the number of products available within package sizes since roll counts are discrete. Dashed lines indicate the 25th and 75th percentiles of the average days' supply purchased by households.

Model Setup

I model a household's purchase decision using a static discrete choice framework. When making a purchase, households consider the price, unit price, quality, quantity, and size of each package and choose the package that maximizes their utility. These features are captured in the household i's indirect utility function:

$$\begin{aligned} U_{ijt} =& \beta_1 Price_{jt} + \beta_2 UnitPrice_{jt} + \beta_3 UnitPrice_{jt} \times Reg_i + \\ & \beta_4 \log(Days_j) + \beta_5 BigPack_j + \beta_6 BigPack_j \times House_i + \\ & \beta_7 SmallPack_j + \beta_8 SmallPack_j \times House_i + \theta_{b(j)} + \epsilon_{ijt}, \end{aligned} \tag{1.13}$$

where $Price_{jt}$ is the total price of product j at time t. Reg_i is an indicator for whether unit price regulations are in effect for household i. $Days_j$ is the number of days the package will last (a function of the number of total sheets in the package and the number of people in the household). $UnitPrice_{jt}$ is the per-day, per-person price of the package, since the yield of a package is how many days it will last. $BigPack_j$ is a dummy for the package having more than 12 rolls and $SmallPack_j$ is a dummy for less than 12 rolls.[39] $House_i$ is an indicator for whether the household lives in a single-family home, with the alternative being an apartment. Finally, $\theta_{b(j)}$ is a brand fixed effect. Brand fixed effects capture quality differences between products. I assume ϵ_{ijt} is iid Type 1 extreme value.

This simple model incorporates the key features necessary to quantify the contribution of cognitive and storage costs to the bulk-buying gap. Preferences for package size (a measure of storage costs) are captured by β_5, β_6, β_7, and β_8, while the effect of displaying per-unit prices is captured by β_3.

The price coefficient is identified using price variation across shopping trips due to shopping at different stores or sales. The size coefficient is identified by variation in the product "concentration" as illustrated in Figure 1.9. That is, given their preferred days' supply (x-value), some households choose large packages and some choose small packages (y-value).

Given these assumptions and the structure of the error term, the probability that household i chooses product j on trip t has a closed form:

$$P_{ijt} = \frac{e^{\beta' x_{ijt}}}{\sum_j e^{\beta' x_{ijt}}}, \tag{1.14}$$

As a result, the log-likelihood function can be written as:

$$LL(\beta) = \sum_t \sum_i \sum_j y_{ijt} \log(P_{ijt}), \tag{1.15}$$

where y indicates whether household i chose product j on shopping trip t. The preference parameters β can then be estimated using maximum likelihood.

Random Coefficients Estimation

The logit model outlined above provides an important starting point, but it cannot capture possible variation in tastes between households. Households may weigh unit

[39] Households bunch at 12-roll packages, so this allows for different package preferences around this bunching point.

prices or package sizes differently based on unobserved factors outside of whether they live in a single-family home or what their regulatory environment is. If this unobserved variation is significant, then errors will be correlated and the estimates from the logit model will be biased.

Furthermore, the Nielsen data has a panel structure and purchases made by the same household are likely to be correlated. The logit specification cannot accommodate kind of correlation, but a random-coefficients specification can. I assume that the unit price, log days' supply, and package size (large and small) coefficients are normally distributed and allow for them to be correlated. I estimate the model using simulated maximum likelihood.[40] To increase accuracy and reduce computational burden, I use pseudo-random Halton draws in the estimation procedure (Hensher and Greene 2003).

1.6 Estimation Results

I estimate this model separately for each income quartile using household purchases from 2016. I observe about 45,500 toilet paper purchases across about 14,800 households at grocery stores and mass merchandisers.

Conditional Logit

Table 1.10 reports model estimates for the conditional logit specification.

The estimation results show that both the price and unit price coefficients are negative, implying that all else equal, households prefer lower prices. Lower income households are more price sensitive than high-income households. The interaction terms reveal that when unit prices are posted, all households are more sensitive to unit prices. This pattern supports the assertion that households respond to the provision of new price information. All households prefer to have more days' supply of toilet paper compared to less. In terms of storage costs, all households have a similar preference against large sizes and, with the exception of the highest-income households, this preference is not significantly different based on their housing type. The highest-income households that live in single-family homes have a much smaller

[40] I use the `mlogit` package which implements Ken Train's Matlab code in R (Revelt and Train 1998).

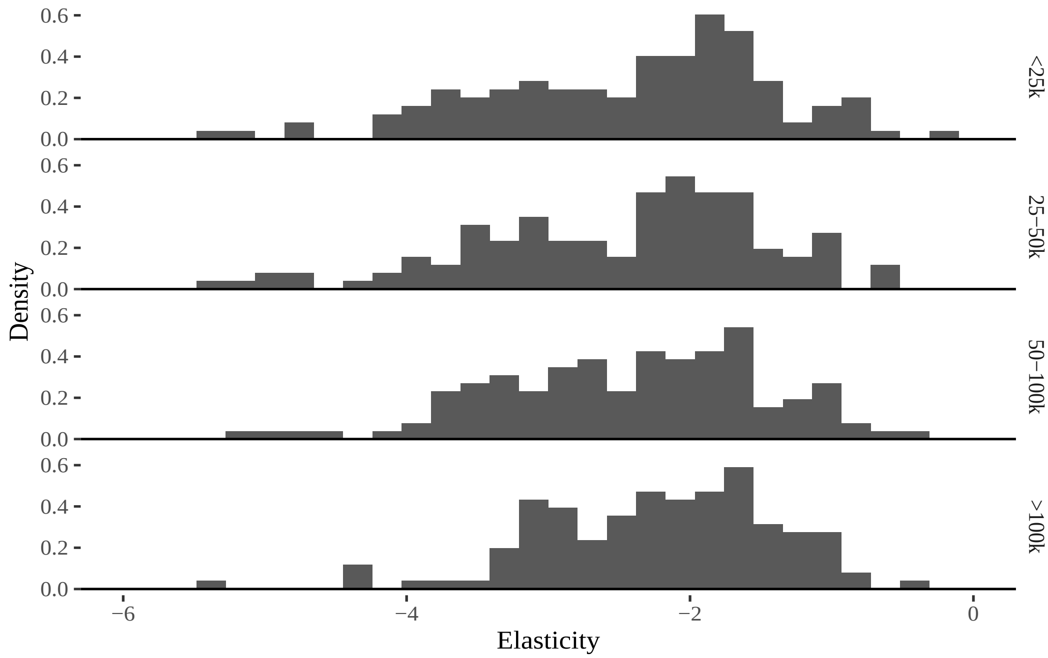
Table 1.10: Multinomial Logit Estimation Results (2016)

	<25k	25-50k	50-100k	>100k
	(1)	(2)	(3)	(4)
Total Price	-0.230^{***}	-0.219^{***}	-0.205^{***}	-0.170^{***}
	(0.009)	(0.005)	(0.004)	(0.006)
Unit Price	-2.173^{***}	-1.919^{***}	-1.946^{***}	-1.902^{***}
	(0.172)	(0.085)	(0.068)	(0.103)
. : Reg	-0.585^{**}	-1.145^{***}	-0.653^{***}	-0.247^{**}
	(0.235)	(0.119)	(0.083)	(0.104)
Log(Days)	0.641^{***}	0.719^{***}	0.694^{***}	0.648^{***}
	(0.048)	(0.031)	(0.029)	(0.048)
Large Size	-0.516^{***}	-0.577^{***}	-0.336^{***}	-0.541^{***}
	(0.096)	(0.070)	(0.067)	(0.126)
. : Home	0.133	0.075	−0.029	0.298^{**}
	(0.113)	(0.078)	(0.072)	(0.129)
Small Size	-0.259^{***}	-0.314^{***}	-0.324^{***}	-0.481^{***}
	(0.057)	(0.044)	(0.048)	(0.091)
. : Home	-0.188^{***}	-0.090^{*}	-0.162^{***}	−0.056
	(0.065)	(0.047)	(0.049)	(0.093)
Brand FE's	Y	Y	Y	Y
Observations	4,968	12,950	17,875	7,942
Log Likelihood	−15,858.260	−40,901.380	−56,134.410	−24,815.100

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Using 2016 Nielsen Consumer Panel and Retail Scanner data, this table presents MLE estimates from Equation 1.13. “Total Price” denotes the total price of the package while“unit price” is the price per day that the package will last (assuming constant consumption of 57 sheets per day (Jaffe 2007)).“Reg” indicates whether unit price regulations are in effect.“Large” indicates packages that are larger than 12 rolls and“small” indicates packages that are smaller than 12 rolls. A 12-roll package is the reference group.“House” indicates if the household lives in a single-family home (reference group is an apartment).

Figure 1.10: Distribution of Price Elasticity by Household Income



Note: Using 2016 Nielsen Consumer Panel and Retail Scanner data, this figure plots the distribution of price elasticities resulting from the estimation of Equation 1.13.

preference against large packages than other income groups. On the other hand, all households also dislike small packages. Under a pure storage costs story, the small packages would have been expected to have a positive sign for low-income households. However, as mentioned in the model specification section, there is bunching at 12-roll packages across households of all types, so this negative sign on the small size is likely a result of that bunching.

Figure 1.10 plots the distribution of own-price elasticities for each product. The majority of elasticities fall between -1.2 and -3.9 with poorer households having slightly larger elasticities (in magnitude).[41]

These results support my earlier findings that unit price regulations affect a household's bulk buying decision and that high-income households have lower storage

[41]Table 4 of Cohen (2008) reports elasticities ranging from -1.94 to -2.54 for paper towels. My estimates cover this range, but are generally much lower with a large mass between -1.2 and -2.0. Demand for toilet paper is likely less elastic than paper towels since cloth kitchen towels or paper napkins can substitute for paper towels. Toilet paper does not have any similar, readily available, substitutes.

costs. In the counterfactuals, I predict how the bulk buying gap changes in response to regulatory changes and reductions in storage costs.

Model Fit

In order to examine how well the model fits the data, I determine if it reasonably predicts the amount of toilet paper that households purchase. To calculate this, I compute the predicted market share for each product and then calculate the expected days' supply purchased based on those market shares. Table 1.11 compares the overall model predictions to the actual data. The model fits the data quite well, even given its parsimony. It slightly over-predicts purchase amounts, but this is primarily because it does not capture some products that are disproportionately popular (or unpopular) relative to what would be expected based on their characteristics. For example, a particular Charmin 6-pack has an 8–10% share for each income group, but based on its characteristics, the model only predicts a 6–7% share. Replacing the brand fixed effect with a product fixed effect would ensure a better fit, but at the cost of reducing the interpretability and intuition of the model. I opt to maintain the parsimony and interpretability of the model and simulate counterfactuals using this specification.

Table 1.11: Multinomial Logit Model Fit (Days' Supply Purchased)

Income	Data	Model
$<$25k	48.64	49.09
25-50k	49.07	49.45
50-100k	51.23	52.34
$>$100k	53.85	54.14

Note: Using 2016 Nielsen Consumer Panel and Retail Scanner data, this table compares the average days' supply of toilet paper purchased in the data with the predicted purchase from the model. I assume an average daily consumption rate of 57 two-ply sheets per day (Jaffe 2007).

Random Coefficients

Table 1.12 reports model estimates for the random coefficients specification.

Table 1.12: Random Coefficient Estimation Results (2016)

	<25k	25-50k	50-100k	>100k
	(1)	(2)	(3)	(4)
Total Price	−0.300***	−0.287***	−0.284***	−0.196***
	(0.011)	(0.007)	(0.006)	(0.009)
Unit Price	−4.114***	−2.642***	−2.969***	−3.243***
	(0.290)	(0.124)	(0.105)	(0.155)
. : Reg	−1.914***	−1.759***	−0.819***	−0.503***
	(0.330)	(0.161)	(0.123)	(0.149)
Log(Days)	1.146***	1.381***	1.452***	0.843***
	(0.078)	(0.050)	(0.047)	(0.071)
Large Size	−1.654***	−1.194***	−1.288***	−1.104***
	(0.174)	(0.096)	(0.098)	(0.181)
. : Home	0.273*	0.101	−0.016	0.422**
	(0.147)	(0.097)	(0.093)	(0.179)
Small Size	−0.437***	−0.345***	−0.504***	−0.807***
	(0.083)	(0.062)	(0.066)	(0.139)
. : Home	−0.173*	−0.241***	−0.221***	−0.202
	(0.095)	(0.067)	(0.069)	(0.140)
$\sigma_{unitPrice}$	6.009***	4.973***	4.791***	3.849***
	(0.241)	(0.116)	(0.092)	(0.113)
$\sigma_{log(Days)}$	1.620***	1.499***	1.812***	1.651***
	(0.062)	(0.039)	(0.035)	(0.056)
σ_{Large}	1.847***	1.327***	1.647***	1.846***
	(0.142)	(0.073)	(0.065)	(0.085)
σ_{Small}	2.111***	2.181***	2.011***	2.353***
	(0.083)	(0.048)	(0.042)	(0.067)
Brand FE's	Y	Y	Y	Y
Observations	4,968	12,950	17,875	7,942
Log Likelihood	−14,502.240	−37,679.860	−51,245.220	−22,705.370

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Using 2016 Nielsen Consumer Panel and Retail Scanner data, this table presents MLE estimates from Equation 1.13. "Total Price" denotes the total price of the package while "unit price" is the price per day that the package will last (assuming constant consumption of 57 sheets per day (Jaffe 2007)). "Reg" indicates whether unit price regulations are in effect. "Large" indicates packages that are larger than 12 rolls and "small" indicates packages that are smaller than 12 rolls. A 12-roll package is the reference group. "House" indicates if the household lives in a single-family home (reference group is an apartment).

The random coefficients estimates show the same qualitative patterns as the conditional logit model. However, there are some notable differences. Each of the random coefficients displays substantial heterogeneity.[42] Lower income households are more price sensitive than high-income households. Furthermore, low-income households have a wider range of sensitivities than higher-income households, as implied by the standard deviations of the unit price coefficient. The regulation and package size coefficients reveal that lower income households are more responsive to unit prices when they are posted and have more severe storage costs.

Figure 1.11 plots the distribution of own-price elasticities for each product using the random coefficients estimates. These are calculated by increasing the price (and the corresponding unit price) of a product by 1% and simulating the new market shares for each product. The percent change in the market shares for the product whose price changed is the own-price elasticity. The majority of elasticities fall between -1.7 and -5.1 with poorer households having larger elasticities (in magnitude).

Model Fit

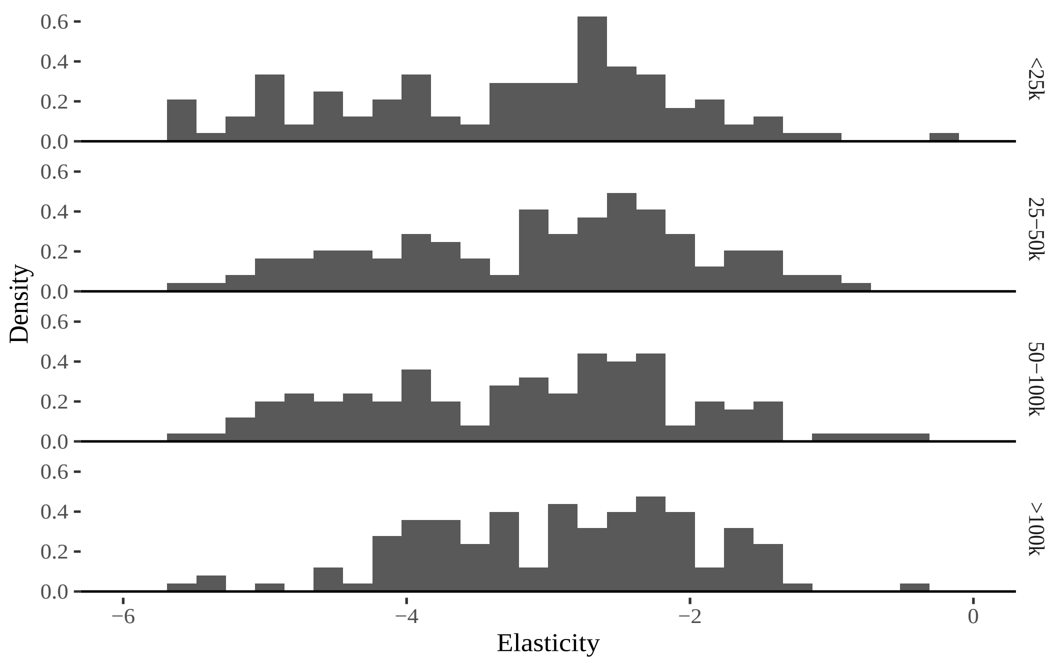
I examine model fit similarly to the conditional logit model. However, since coefficients are random, the choice probabilities take the following form:

$$P_{ijt} = \int \frac{e^{\beta' x_{ijt}}}{\sum_j e^{\beta' x_{ijt}}} f(\beta) d\beta, \tag{1.16}$$

I use simulation to approximate the integral by taking 1,000 draws from the joint distribution of β. Table 1.13 compares the overall model predictions to the actual data. The model over-predicts the amount purchased across all households, compared to the conditional logit model, but this is primarily because it over-predicts the purchases of particularly large generic packages. For example, a particular generic 12-pack has a 1–2% share for each income group, but based on its characteristics, the model predicts a 3–5% share. The model assumes that all generic brands are equal, but in reality, it may be the case that generic brands differ based on the retailer that sells them. This additional dimension of heterogeneity could be captured by more granularly defining brands by the retailer that sells them.

[42] Since each random coefficient was assumed to be normally distributed, some households may

Figure 1.11: Distribution of Price Elasticity by Household Income



Note: Using 2016 Nielsen Consumer Panel and Retail Scanner data, this figure plots the distribution of price elasticities resulting from the estimation of Equation 1.13, using random coefficients.

Table 1.13: Random Coefficient Model Fit (Days' Supply Purchased)

Income	Data	Model
$<$25k	48.64	50.67
25-50k	49.07	52.45
50-100k	51.23	54.48
$>$100k	53.85	54.93

Note: Using 2016 Nielsen Consumer Panel and Retail Scanner data, this table compares the average days' supply of toilet paper purchased in the data with the predicted purchase from the model. I assume an average daily consumption rate of 57 two-ply sheets per day (Jaffe 2007).

1.7 Counterfactuals

Using the parameter estimates from the previous section, I predict how households respond to lower storage costs and universal unit price regulation. For these counterfactual exercises, I compare all counterfactual results to a "base case" of predicted purchases given their current shopping environment. I consider two counterfactual scenarios:

1. **Unit-Price Regulation:** Unit-price regulations are adopted everywhere.
2. **Reduced Storage Costs:** All households have the same storage costs (i.e., size preferences) as high-income households.

For the unit-price regulation scenario, I set each household's unit price coefficient equal to the sum of its coefficient and the regulation interaction term. For households making under \$25,000, their unit price coefficient becomes $-2.173 - 0.585 = -2.758$. For the reduced storage cost scenario, I set all size coefficients equal to the coefficients for households making over \$100,000.

Table 1.14 reports the counterfactual predictions for both the conditional logit (top panel) and the random coefficients model (bottom panel). I will discuss the conditional logit results first. When unit price regulations are adopted, all households increase their purchase quantity, with middle-income groups increasing the most. Compared to the original days' supply purchased by high-income households, low-income households close the gap by 16% from 5.05 days' supply to 4.25 days' supply after unit price regulations are universally adopted. Equalizing storage costs reduces the gap by another 1.51 days' supply, an additional 30% of the gap. After making these two changes, all but the lowest-income households buy almost the same amount as the highest-income households. The remaining difference is due to brand preferences, particularly because low-income households are more willing to buy generic brands and less willing to by premium brands.

The random coefficient counterfactuals, while overpredicting the average days' supply purchased, predicts a gap of 4.26 days' supply between high- and low-income

have a non-intuitive valuation for product attributes, such as a positive valuation for unit price. Sign restrictions can be imposed by assuming alternative distributions, such as a log-normal distribution.

Table 1.14: Bulk Purchasing Counterfactual Simulation Results

Conditional Logit			
Income	Base	+ Unit Price Regs	+ Rich Storage
<25k	49.09	49.89	51.40
25-50k	49.45	51.40	53.88
50-100k	52.34	53.69	54.11
>100k	54.14	54.72	54.72
Random Coefficients			
Income	Base	+ Unit Price Regs	+ Rich Storage
<25k	50.67	52.67	56.86
25-50k	52.45	55.23	58.11
50-100k	54.48	55.91	58.47
>100k	54.93	55.83	55.83

Note: This table reports predicted package quantities purchased by households using model estimates of Equation (1.13). Units are number of days the chosen package will last assuming average daily consumption rate of 57 two-ply sheets (Jaffe 2007). The "Unit Price Regs" scenario imposes unit price regulations everywhere. The "Rich Storage" scenario imposes that all households have the same preferences for "large" and "small" packages as households making over $100k. Scenarios are cumulative.

households. After universally adopting unit price regulations, all households increase their purchasing, but the gap between high- and low-income households shrinks to 3.16 days' supply. Equalizing storage costs actually reverses the gap with the lowest income households buying 1.03 days' supply more than high-income households. This dramatic increase is driven by the fact that middle- and low-income households continue to be highly price-sensitive and have a stronger preference for large quantities.

These counterfactuals support the main findings from Section 1.4 which showed that unit price regulations increase bulk buying and that storage costs are a substantial factor preventing households from buying in bulk. Since the earlier sections examined bulk purchasing across all non-food products, I repeat the earlier analysis on mover households specifically for toilet paper purchases. I estimate a modified version of Equation 1.5 which replaces share of bulk purchases with log days' supply of toilet paper:

$$Log(DaysSupply)_{it} = \alpha + \beta_1 Reg_{it} + \beta_2 SingleFamily_{it} + \gamma X_{it} + \lambda_i + \lambda_t + \epsilon_{it}, \quad (1.17)$$

Table 1.15 shows that households increase the days' supply purchased by 3.5% when unit prices are posted and by 2.6% when they move into a single-family home. The model predictions are in line with these changes. The conditional logit model predicts that purchasing increases by 1.1–3.9% when unit prices are posted (compared to 3.5% above) and by 0.8–4.8% when storage costs are reduced (compared to 2.6% above). The random coefficients model predicts that purchasing increases by 1.6–5.3% when unit prices are posted (compared to 3.5% above) and by 4.6–8.0% when storage costs are reduced (compared to 2.6% above). The reduced-form estimates and model predictions line up quite well with regards to posting unit prices. However, there are some differences between the two types of estimates with respect to storage costs. Part of this difference may be because home type (apartment or single-family home) does not capture true storage costs while the random coefficients model offers more flexibility to capture heterogeneity between households even with the same type of housing. For example, the presence of a basement, garage, or even the number of bathrooms may all influence the storage costs for toilet paper, but those are all differences that can exist within single-family homes.

Overall, reducing cognitive costs and increasing the salience of unit prices helps households make better value decisions, and generate a strong boost to bulk buying. Adopting unit price regulations are a relatively straightforward policy approach to encourage bulk buying, especially compared to the challenge of feasibly reducing storage costs for low-income households.

Table 1.15: Effect of Unit Price Regulations and Housing Changes on Toilet Paper Purchases

	(1)	(2)
Regulation	0.029^{*}	0.035^{**}
	(0.015)	(0.015)
Single-Family Home		0.026^{***}
		(0.006)
Household FE	Y	Y
Year FE	Y	Y
Demographics	N	Y
Observations	4,553,957	4,553,957
Adjusted R^2	0.507	0.508
Note:	*p<0.1; **p<0.05; ***p<0.01	

Note: Using 2004–2017 Nielsen Consumer Panel data and state-level regulations, this table shows estimates of Equation 1.5 which regresses household bulk buying on unit price regulation after controlling for household fixed effects and changes in household characteristics. "Regulation" denotes the estimated effect of moving from a state without regulation to a state with regulation. "Single-Family Home" indicates that household lives in a single-family home with the reference category being an apartment. Standard errors are clustered at the household level.

1.8 Conclusion

This paper documents the new fact that low-income households are less likely to take advantage of quantity discounts relative to high-income households. This gap is especially large for storable, necessity items like toilet paper and paper towels. If low-income households bought in bulk like high-income households, they could save 5% on grocery items, saving an aggregate of $5.4 billion annually. I provide evidence that *cognitive costs*, *store preferences*, *budget constraints*, and *storage costs* contribute to this gap.

By using state-level variation in whether or not retailers have to display unit prices, I find that displaying unit prices reduces cognitive costs and increases bulk buying. Then, I show that *where* a household shops accounts for a large portion of this disparity and that warehouse clubs increase bulk buying, but only for middle- and high-income households. Low-income households are unlikely to shop at warehouse clubs, even if they are nearby. Next, I demonstrate that low-income households slightly decrease bulk buying towards the end of the month, when budgets are tighter. Finally, I show that households increase bulk buying when they move to larger housing, supporting the fact that storage costs also influence the bulk buying decision.

Combining these features into a discrete choice model of toilet paper purchases, I predict how households' bulk purchasing changes if unit-price regulations are adopted universally and if storage costs are reduced. I find that posting unit prices closes the bulk buying gap by 26% and reducing storage costs completely closes the gap with middle- and low-income households buying larger quantities than high-income households.

This paper is one of the first to focus on consumer's take-up of quantity discounts and explore the factors that contribute to this decision. It provides evidence that *cognitive costs*, *store preferences*, *budget constraints*, and *storage costs* affect a household's bulk buying decision. These differences have substantial financial consequences for the poorest households and are likely to generate systematic underestimates of consumption inequality if quantity discounts offset quality differences between products. Additionally, if the prices of large and small packages evolve differently, then households may experience substantial changes in their buying power. Future work will determine the extent to which inequality and inflation measures are underestimated because of quantity discounts.

References

Aguiar, Mark and Erik Hurst (2005). "Consumption versus expenditure". In: *Journal of political Economy* 113.5, pp. 919–948.

— (2007). "Life-cycle prices and production". In: *American Economic Review* 97.5, pp. 1533–1559.

Allcott, Hunt et al. (2019). "Food deserts and the causes of nutritional inequality". In: *The Quarterly Journal of Economics* 134.4, pp. 1793–1844.

Bauner, Christoph and Emily Wang (2019). "The Effect of Competition on Pricing and Product Positioning: Evidence from Wholesale Club Entry". In: *International Journal of Industrial Organization*.

BLS (Apr. 2019). *Consumer Expenditures in 2017*. Tech. rep. 1080. Bureau of Labor Statistics. URL: https://www.bls.gov/opub/reports/consumer-expenditures/2017/home.htm.

Bogomolova, Svetlana and Ian Jarratt (2016). "Unit pricing in supermarkets: Review of past evidence from academic and industry studies". In: *Available at SSRN 2853977*.

Chetty, Raj, Adam Looney, and Kory Kroft (2009). "Salience and taxation: Theory and evidence". In: *American economic review* 99.4, pp. 1145–77.

Chung, Chanjin and Samuel L Myers Jr (1999). "Do the poor pay more for food? An analysis of grocery store availability and food price disparities". In: *Journal of consumer affairs* 33.2, pp. 276–296.

Cohen, Andrew (2008). "Package size and price discrimination in the paper towel market". In: *International Journal of Industrial Organization* 26.2, pp. 502–516.

Coibion, Olivier, Yuriy Gorodnichenko, and Dmitri Koustas (2017). *Consumption inequality and the frequency of purchases*. Tech. rep. National Bureau of Economic Research Working Paper No. 23357.

DellaVigna, Stefano and Matthew Gentzkow (2019). "Uniform pricing in us retail chains". In: *The Quarterly Journal of Economics* 134.4, pp. 2011–2084.

Federal Highway Administration (2017). *2017 National Household Travel Survey*. Tech. rep. U.S. Department of Transportation. URL: https://nhts.ornl.gov.

General Accounting Office (1975). *Food Labeling: Goals, Shortcomings, and Proposed Changes*. Tech. rep. General Accounting Office. URL: https://www.gao.gov/assets/120/115505.pdf.

Griffith, Rachel et al. (2009). "Consumer shopping behavior: how much do consumers save?" In: *Journal of Economic Perspectives* 23.2, pp. 99–120.

Hendel, Igal and Aviv Nevo (2006). "Measuring the implications of sales and consumer inventory behavior". In: *Econometrica* 74.6, pp. 1637–1673.

Hensher, David A and William H Greene (2003). "The mixed logit model: the state of practice". In: *Transportation* 30.2, pp. 133–176.

Jaffe, Chuck (Oct. 11, 2007). "On a roll". In: *MarketWatch*. URL: https://www.marketwatch.com/story/on-a-roll-toilet-paper-purchases-set-pace-for-weightier-matters.

Jarmin, Ronald S, Shawn D Klimek, and Javier Miranda (2009). "The role of retail chains: National, regional and industry results". In: *Producer dynamics: New evidence from micro data*. University of Chicago Press, pp. 237–262.

Kunreuther, Howard (1973). "Why the poor may pay more for food: theoretical and empirical evidence". In: *The journal of Business* 46.3, pp. 368–383.

Mitchell, Vincent-Wayne, David Lennard, and Peter McGoldrick (2003). "Consumer awareness, understanding and usage of unit pricing". In: *British Journal of Management* 14.2, pp. 173–187.

Miyazaki, Anthony D, David E Sprott, and Kenneth C Manning (2000). "Unit prices on retail shelf labels: an assessment of information prominence". In: *Journal of retailing* 76.1, pp. 93–112.

NIST (2019). *Uniform Laws and Regulations in the Areas of Legal Metrology and Fuel Quality (Handbook 130)*. Tech. rep.

Orhun, A Yeşim and Mike Palazzolo (2019). "Frugality Is Hard to Afford". In: *Journal of Marketing Research* 56.1, pp. 1–17.

Revelt, David and Kenneth Train (1998). "Mixed logit with repeated choices: households' choices of appliance efficiency level". In: *Review of economics and statistics* 80.4, pp. 647–657.

Rose, Veronica (Nov. 2000). *Unit Pricing and Electronic Shelf Labels*. Tech. rep. 2000-R-1044. Connecticut Office of Legislative Research. URL: https://www.cga.ct.gov/2000/rpt/2000-R-1044.htm.

Sprott, David E, Kenneth C Manning, and Anthony D Miyazaki (2003). "Grocery price setting and quantity surcharges". In: *Journal of Marketing* 67.3, pp. 34–46.

Statista (2019). "U.S. Population: Do you use toilet paper?" In: URL: https://www.statista.com/statistics/275965/us-households-use-of-toilet-paper/.

Stephens Jr, Melvin (2003). "" 3rd of tha Month": Do Social Security Recipients Smooth Consumption Between Checks?" In: *American Economic Review* 93.1, pp. 406–422.

Talukdar, Debabrata (2008). "Cost of being poor: retail price and consumer price search differences across inner-city and suburban neighborhoods". In: *Journal of Consumer Research* 35.3, pp. 457–471.

USDA (2019). *Supplemental Nutrition Assistance Program Participation and Costs*. Tech. rep. URL: https://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap.

1.A Appendix

Data Appendix

The Nielsen Consumer Panel consists of about 40,000–60,000 US households that provide information on their shopping purchases using in-home scanners or Nielsen's mobile app. Panelists are geographically dispersed and demographically balanced. Households are recruited based on key demographic characteristics, primarily household size, income, age, education, presence of children, race, ethnicity, and occupation. To generate national averages, Nielsen assigns each household a projection factor.

Households are recruited through direct mail and online invitations. To incentivize households to remain in the panel, Nielsen provides monthly prize drawings, sweepstakes, points, and regular communication and support to panelists. Nielsen tries to ensure that incentive methods are non-biasing and regularly tests for its correlation with retention rates. To ensure data quality, Nielsen filters out any households that are poor reporters and do not meet minimum spending thresholds based on their household size. All households in the sample meet this threshold for the full year.

Demographic variables are recorded and updated annually. For my analysis, I collapse some of the demographic variables into more aggregate categories. Household composition measures the number adults and children residing in the home. Marital status is an indicator for whether the head of household is married or not (I do not distinguish between single, divorced, or widowed). Education is an indicator for whether at least one head of household completed college. Housing variables indicate whether a household lives in a single-family home, and apartment, or a mobile home. Finally, age is the age of the head of household. In the case of two heads, I average the two ages.

To construct my analysis sample, I remove any households where the head of household is a student or a member of the military because these households likely have different living arrangements that are not representative of a typical household's decision (i.e. on campus housing or barracks are different than traditional homes and apartments). I drop any households living in mobile homes as well because this type of housing could include a wide range of house types including RVs and manufactured homes. I also remove any households making less that $5,000 and those that could not be geocoded based on their ZIP code.[43] Finally, some households were dropped because they could not be matched to tract-level vehicle access data.[44] Table 1.16 reports how many households were removed based on this cleaning procedure.

Table 1.16: Homescan Sample Construction

Step	HH
Starting HH:	178,232
Exclude military and students:	175,102
Exclude Households under 5k:	174,106
Exclude Mobile Homes:	167,065
Drop ZIPs Not Geocoded:	166,366
Cannot Be Matched to Car Access:	166,164

Note: Using 2004–2017 Nielsen Consumer Panel data, this table reports the number of unique households in the sample after each step of data refinement.

In the purchase data, I exclude alcohol, tobacco, pet items, health and beauty items, general merchandise, "magnet," and "deferred" product categories from my analysis. Alcohol and tobacco are excluded because of their addictive qualities, which may induce peculiar purchase patterns. For example, a smoker may choose to only buy one pack of cigarettes with the intention of quitting even though a full carton may deliver a better value. Pet items are excluded to focus on products intended for human consumption. I exclude health and beauty items and general merchandise because products such as trash cans, printers, eye shadow, and antacids are unlikely to be bought in bulk or have irregular consumption patterns. "Deferred" categories are categories that Nielsen has stopped tracking, so to maintain a consistent sample

[43]I use the 2017 Census Gazetteer to assign ZIP codes to the latitude and longitude of their population-weighted centroid

[44]Vehicle access data comes from the 2009–2013 American Community Survey, which asks how individuals get to work. There is limited variation in this measure since most respondents have vehicle access. For context, only 4% of Nielsen households live in Census tracts less than 90% access to cars.

of products, these are excluded from my analysis. Finally, "magnet" purchases are items which do not have a UPC code, such as fresh fruits and vegetables, deli counter items, or bakery items. Because these items are only recorded for a subset of Nielsen households and are not standardized, I also exclude them from my analysis. This process leaves me with 721 unique product categories. Because this paper focuses on bulk purchases, I also exclude 28 categories that have five or fewer sizes across all possible products.[45] Overall, the products analyzed are common household staples including almost all food categories, basic toiletry items, and non-food essentials like toilet paper, soaps/detergents, and diapers. See Table 1.17 for summary statistics of the top 20 product categories by annual spending.

To compare sizes across different product categories, I assign each product to its quintile in the size distribution for that product category. I assign quintiles based upon the sample quintiles of product sizes to ensure that each quintile has 20% of available products in its support. An alternative strategy would assign quintiles based on cutting the range of product sizes into equal intervals. However, in some product categories, this risks generating quintiles with sparse support when there is an especially large package available. As an example, consider eggs. Most packages contain 6, 12, or 18 eggs, but there are some products that offer up to 15-dozen eggs (180 eggs). Generating quintiles by cutting the available range into equal intervals would generate quintiles of 1-36, 37-72, 73-108, 109-144, 145-180 which would assign almost all packages to the first quintile and the fifth quintile. Using the sample quintiles generates a more even distribution ensuring better support of each quintile. For products with a narrow range of sizes that fall in multiple quintiles, I assign the product to the minimum quintile. For example, over 60% of egg products are dozens, which covers three quintiles. I assign all products with 12 or fewer eggs to the first quintile.

[45] These excluded categories are: jelled aspic salad, sour cream sauce mix, canned roast beef, canned roast beef hash, retort pouch bags, prepared sandwiches, canned rice, canned dumplings, canned bread, frozen vegetables in pastry, frozen grapefruit juice, frozen grape juice, frozen orange juice, frozen cream substitutes, canned ham patties, bathroom accessory, packaged soap, borateem, dry starch, grease relief, bathroom brushes, miscellaneous brushes, thermometers, dustpans, feather dusters, laundry baskets, sanitary belts, gift package with candy or gum.

Table 1.17: Summary Statistics of Top 20 Product Categories in Nielsen Data (2017)

Product	Annual Spending	SD	Avg. Price	SD	Avg. Size	SD
Soft Drinks	79.38	139.02	4.75	4.17	85.87	53.81
Diet Soft Drinks	74.82	132.79	4.73	5.07	84.18	65.06
Milk	65.65	77.02	3.11	1.79	97.79	35.00
Cereal	57.97	68.37	4.06	2.10	18.05	8.17
Toilet Paper	56.15	49.47	11.44	7.09	17.09	10.51
Yogurt	55.00	75.68	3.28	2.17	17.25	15.22
Coffee	53.97	61.69	8.60	5.74	21.84	11.05
Bread	50.03	47.09	2.88	1.52	20.54	4.64
Cookies	46.97	57.60	3.59	3.44	13.02	6.39
Fresh Meat	46.96	62.86	7.75	5.03	30.48	24.97
Frozen Pizza	44.48	60.64	5.99	3.67	20.69	12.48
Bottled Water	44.06	73.46	4.21	3.75	261.91	181.39
Fresh Fruit	42.68	64.91	4.28	2.06	1.93	1.31
Chocolate Candy	41.05	53.83	3.91	3.67	8.64	9.15
Detergent	40.17	45.29	10.05	7.85	99.52	61.23
Shredded Cheese	39.16	42.80	4.21	2.45	13.37	10.98
Bacon	37.63	45.44	6.87	4.67	17.42	11.88
Ice Cream	37.36	50.34	4.43	2.03	46.80	24.47
Potato Chips	35.99	41.71	3.04	1.89	8.87	3.81
Canned Soup	32.39	38.36	3.21	2.22	22.07	17.33

Note: Using 2004–2017 Nielsen Consumer Panel data, this table reports summary statistics for the top 20 product categories by total spending. Annual spending is the average spending in that product category among households that purchased in that product category over the course of the year. Average price and average size are the average prices and sizes of products purchased in their corresponding category. All estimates are weighted using Nielsen's projection weights. Prices are in nominal 2017 dollars. Sizes are reported in common units for for that category (e.g. ounces for milk).

Quantity Discounts and Coupon Savings

This section compares savings from quantity discounts to savings from coupons. To be conservative, I compare the savings from redeemed coupons (likely higher than the average savings of all coupons offered) to savings offered by quantity discounts (likely lower than quantity discounts actually redeemed). For each product purchased in the Consumer Panel data, households can input the value saved if they used a coupon. For each product category, I compute the average discount across all products in that category.

I then estimate quantity discount savings based on moving from a product in the second quintile to the fourth quintile of the size distribution. This leaves out small product sizes that may have high unit prices due to convenience (e.g., a 20-oz soda bottle at the checkout counter) and especially large sizes that may not be widely available at all stores. This range covers sizes that households are likely to consider when making their purchase decision.

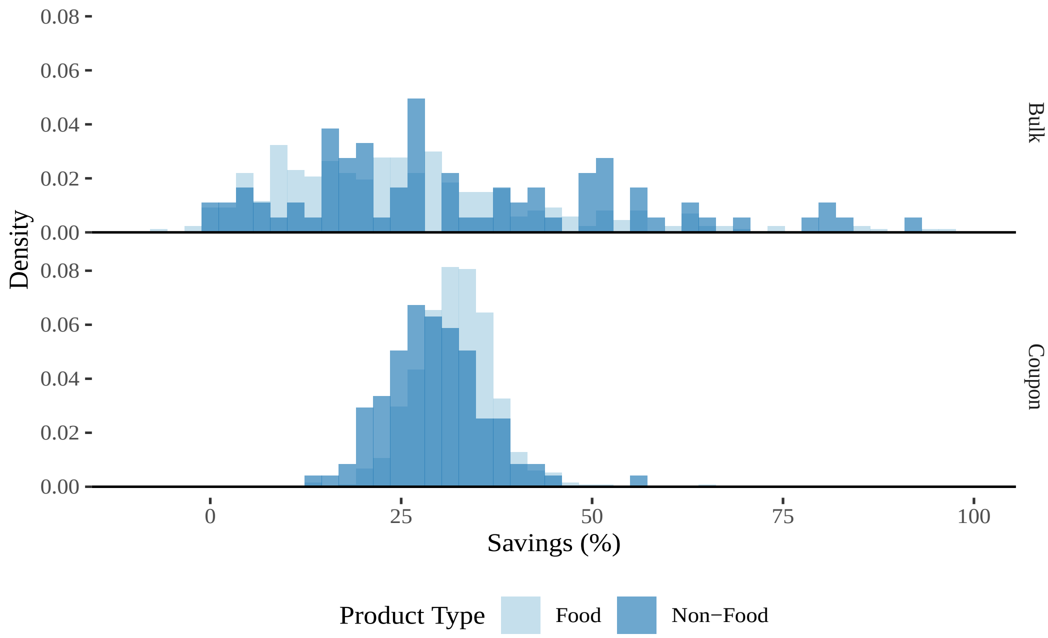
Figure 1.12 plots the distribution of coupon savings and estimated bulk savings for food and non-food products. Coupon savings are narrowly clustered with a median savings of 31% for non-food products and 33% for food products. Bulk discounts have lower median savings for non-food and food products of 27% and 23%, respectively, but are more widely dispersed, even exceeding 50% savings for some non-food products.[46] Coupon savings are similar across product categories while there is substantial variation in quantity discounts with non-food products offering higher savings.

Bulk Buying Across Popular Categories

Across popular spending categories, these gaps are particularly large in storable, non-food categories like paper towels and toilet paper, where households making over $100,000 are more than twice as likely to buy in bulk compared to households making under $25,000. In popular food categories like milk and eggs, there is little relationship or even a negative relationship between income and bulk buying (See Figure 1.13).

[46] Smaller shifts, such as from the second to third quintile or third to fourth quintile generate smaller savings, but still preserve the long right tail primarily for non-food products.

Figure 1.12: Estimated Savings from Coupons and Bulk Discounts



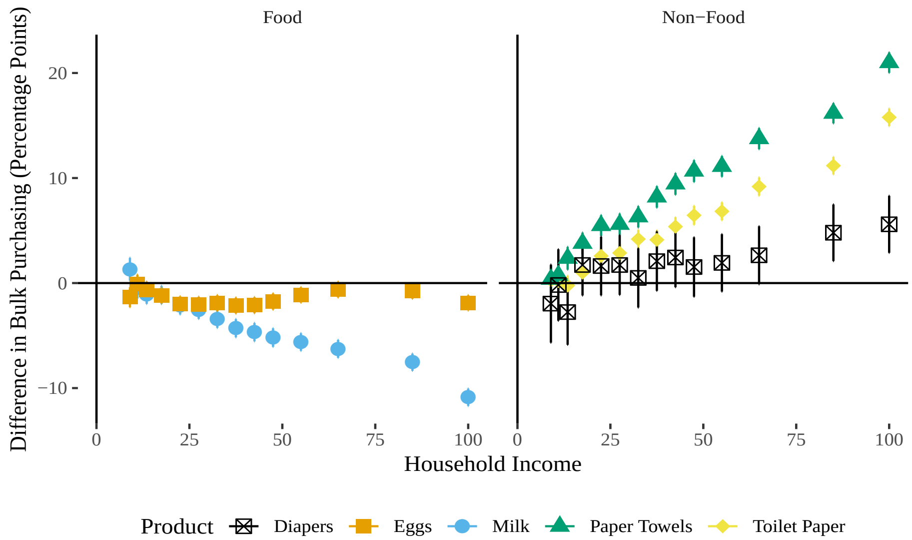
Note: Using 2004–2017 Consumer Panel and 2016 Retail Scanner data, this figure plots the distribution of savings from coupons and quantity discounts. For each coupon redemption, the percent savings are the ratio of the coupon value to the product's price. These savings are then averaged across all purchases in that product category. Bulk discounts are computed using coefficient estimates obtained from Equation (1.1) relating log unit prices to log package sizes. Bulk savings are the estimated savings obtained from moving from the second to the fourth quintile of the size distribution for each product category.

Alternative Calculation of Missed Quantity Discounts

An alternative way of calculating savings from quantity discounts is to calculate first-best savings obtained from purchasing the lowest unit-priced item available, since even high-income households may not be buying at the lowest unit price. I compute this by taking the difference between the unit price paid by each household and the lowest unit price available in the store, given the household's brand preference. I get this information through linking the Nielsen Consumer Panel with the Nielsen Retail Scanner data.

I compute the first-best savings a household could obtain for toilet paper, diapers, milk, and eggs using the following approach. First, for each shopping trip, I compute the lowest unit price the household could have paid given its brand and store choice

Figure 1.13: Bulk Purchasing by Household Income (Selected Products)



Note: Using 2004–2017 Nielsen Consumer Panel data, this figure plots the income bin coefficients from Equation (1.2), which regresses the share of annual purchases that were bulk packages on household characteristics as well as market and year fixed effects. This regression is estimated for milk, eggs, diapers, toilet paper, and paper towels. Nielsen projection weights are used to ensure national representativeness. Households making \$5–8k are the reference group. Standard errors are clustered at the DMA level.

in that week. The difference in unit prices relative to the unit price chosen is a household's first-best savings for that purchase. Then, to get the average savings for a household, I compute the expenditure-weighted average savings across all purchases for each household. Based on this measure, Table 1.18 reports average excess spending by income group, computed for a family of four.

Overall, households could save over 30% by buying in bulk and low-income households could save even more. I estimate the differences in savings between households from the following regression:

$$Y_{imt} = \alpha + \sum_{q} \beta^q Income_{imt} + \gamma X_{imt} + \lambda_{mt} + \epsilon_{imt}, \tag{1.18}$$

where Y_{imt} is the excess spending of household i in market m in year t. $Income_{imt}$ is the household's income bin and X_{imt} consists of household characteristics. λ_{mt} is a market-year fixed effect. Table 1.19 shows that low-income households miss out on 1.7–1.8 percentage points more savings than high-income households and the excess

Table 1.18: First-Best Savings by Household Income and Product

	Non-Perishable		Perishable	
Income	Toilet Paper	Diapers	Milk	Eggs
<\$25k	0.36	0.33	0.31	0.17
\$25-50k	0.35	0.33	0.30	0.17
\$50-100k	0.34	0.33	0.31	0.17
>\$100k	0.33	0.31	0.33	0.18

Note: This table uses 2006–2016 Nielsen Retail Scanner and Consumer Panel data to compute average savings a household could achieve given its brand and store choice. Average savings for a family of four is reported above. For example, a household making <\$25k could save 36% by purchasing at the lowest unit price available.

spending is primarily in non-food categories like toilet paper (36% savings) and diapers (33% savings) as opposed to food categories like milk (31% savings) and eggs (17% savings). Given the perishability of food items, these savings may not be realized if the product perishes before it can be consumed.

Table 1.19: Regression Results of First-Best Savings Across Household Income and Products

	Diapers	Toilet Paper	Eggs	Milk
	(1)	(2)	(3)	(4)
25-50k	-0.010^{**}	-0.005^{***}	0.001	-0.002
	(0.005)	(0.001)	(0.001)	(0.001)
50-100k	-0.015^{***}	-0.013^{***}	0.004^{***}	0.002^{**}
	(0.005)	(0.001)	(0.001)	(0.001)
>100k	-0.018^{***}	-0.017^{***}	0.018^{***}	0.010^{***}
	(0.005)	(0.002)	(0.002)	(0.001)
Demographics	Y	Y	Y	Y
Market-Year FE	Y	Y	Y	Y
Observations	36,903	182,415	194,413	247,451
Adjusted R^2	0.012	0.071	0.117	0.231
Note:	*p<0.1; **p<0.05; ***p<0.01			

Note: This table uses 2006–2016 Nielsen Retail Scanner and Consumer Panel data and reports the income coefficients of Equation (1.18), which regresses savings on household characteristics as well as a market and year fixed effect. Nielsen's projection weights are used for national representativeness.

Overall, low-income households could benefit substantially from buying in bulk and obtaining lower unit prices. Furthermore, these savings are likely to be more important for low-income households since the marginal utility of an additional dollar of savings is likely to be higher than for high-income households. This analysis also provides evidence that all households could benefit from purchasing at the lowest unit price.

Bulk Buying by Store Type or Chain Size

In this section, I analyze whether the effect of unit pricing differs by store type or chain size. Unit price regulations are only at the state level, but retailers are free to post (or not post) unit prices as long as they are within the boundaries of the law. Large chains may post prices uniformly across all stores in a way that meets the strictest requirements they are subject to. On the other hand, regional chains or independent stores may more closely mirror the laws of the state they are located in. I estimate Equation 1.4 using annual household bulk buying at specific stores types or within different chain sizes. Each observation is at the household-year-channel (or chain) level. For example, bulk items accounted for 50% of Household A's grocery store purchases while bulk items accounted for 100% of Household A's warehouse club purchases.

Table 1.20 shows that in the cross-section, stricter unit price regulations are associated with more bulk buying primarily for grocery stores, drug stores, and warehouse clubs. Households in states with strict unit price regulations buy in bulk five percentage points more at grocery stores compared to households in states without any pricing regulations. Since grocery stores tend to be regional or independent, the large positive relationship provides strong evidence that unit price regulations can increase bulk purchasing. Grocery stores also have the richest variety in Nielsen's data with over 900 unique retailers being captured compared to 65 drug stores, 25 discount stores, 17 dollar stores, and 10 warehouse clubs.[47] Other store types exhibit smaller

[47] Nielsen's categorization includes a "catch-all" category that is not unique to a particular retailer, so it actually uniquely captures 64 drug stores and purchases at other drug stores are assigned to the last "catch-all" drug store. Generally, larger retailers are uniquely tracked and smaller ones may fall into the "catch-all" category.

Table 1.20: Unit Price Regulations and Bulk Buying by Store Type

	Grocery	Drug	Discount	Dollar	Warehouse
	(1)	(2)	(3)	(4)	(5)
Vol. Disp	0.011	-0.011^{***}	-0.006	-0.006	-0.004^{*}
	(0.009)	(0.003)	(0.005)	(0.005)	(0.002)
Mand. Disp	0.029^{***}	0.009	0.014^{*}	-0.022^{**}	0.010^{***}
	(0.005)	(0.007)	(0.008)	(0.010)	(0.003)
Mand. Disp, Strict	0.054^{***}	0.018^{***}	0.002	-0.006	0.006^{***}
	(0.009)	(0.003)	(0.004)	(0.006)	(0.002)
Avg Bulk	0.36	0.29	0.49	0.35	0.95
Demographics	Y	Y	Y	Y	Y
Omit California	Y	Y	Y	Y	Y
Observations	618,029	298,166	562,749	328,607	267,759
Adjusted R^2	0.011	0.003	0.005	0.002	0.001

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Using Nielsen 2004–2017 Consumer Panel data combined with state-level regulations, this table shows the results of estimating Equation 1.4. The dependent variable is the annual share of bulk purchases made by households in a particular channel (i.e. store type) and the independent variables are an ordered measure of regulatory stringency. California is omitted because it is the only state that has voluntary unit price, but strict requirements on how unit prices are displayed. Standard errors are clustered by state.

or insignificant effects, which could be because these are generally large chains that have more uniform pricing practices across all locations.

Table 1.21 shows the results by chain size. Following Jarmin, Klimek, and Miranda (2009), I define a "local" chain as only having locations in one state, a "regional" chain has locations in two to ten states, and a "national" chain has locations in more than ten states. In the cross-section, stricter unit price regulations are associated with more bulk buying across all chain types. The effect is strongest for local and regional chains, exhibiting a six to seven percentage point increase in bulk buying relative to states without unit price regulations. National chains still have significant differences, but they are a more moderate three to four percentage point difference relative to states without regulations. Overall, the relative effect is strongest for the smaller chains that are likely to only be subject to a limited set of regulations and the effect is weaker for national chains which may be more likely to adopt pricing practices that satisfy the

Table 1.21: Unit Price Regulations and Bulk Buying by Retailer Size

	Local	Regional	National
	(1)	(2)	(3)
Vol. Disp	0.063***	0.030	0.010
	(0.022)	(0.025)	(0.009)
Mand. Disp	−0.067	−0.018	0.039***
	(0.053)	(0.014)	(0.009)
Mand. Disp, Strict	0.064***	0.070***	0.026***
	(0.016)	(0.014)	(0.006)
Avg Bulk	0.33	0.32	0.49
Demographics	Y	Y	Y
Omit California	Y	Y	Y
Observations	1,578	43,756	668,566
Adjusted R^2	0.008	0.005	0.037
Note:	*p<0.1; **p<0.05; ***p<0.01		

strictest requirements nationwide.

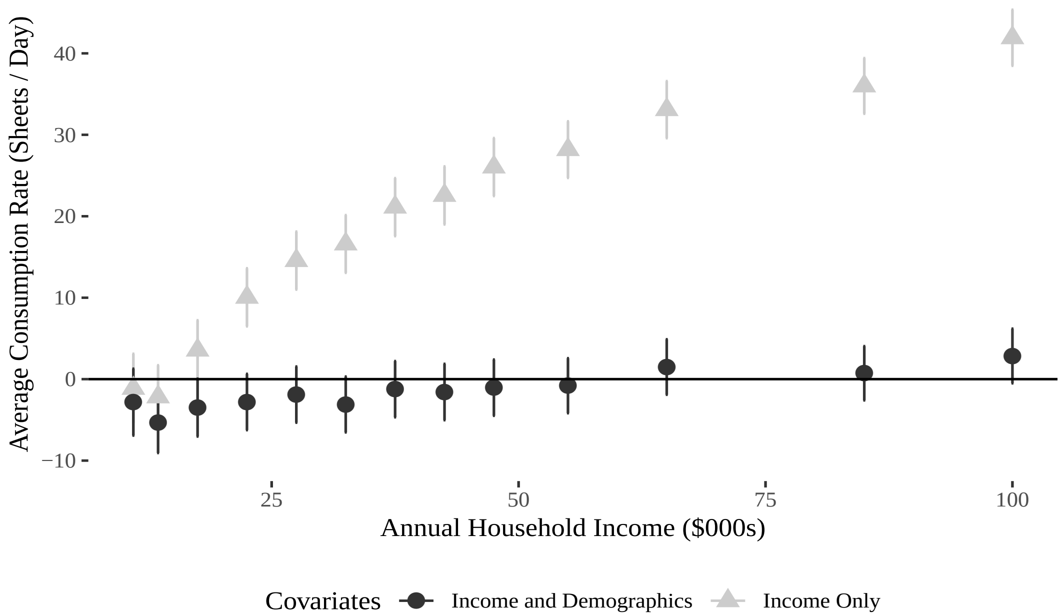
Annual Consumption Analysis

I show that income is not predictive of a household's toilet paper consumption rate first using basic OLS regressions. I then formalize the result using a 100-fold cross-validated elastic net regression to select the most predictive variables. If income and toilet paper consumption are related, then an OLS regression will extract the correlation.

First, I compute a household's daily consumption by aggregating the total number of sheets purchased by a household in a given year, excluding the final purchase of the year since it may not be consumed within the year. I divide this total by the number of days between the first and last purchase of the year to get a household's average daily consumption rate. This method avoids complications where end of the year inventory may be carried over to the following year or a household may start the year with some inventory.

Given a household's average daily consumption rate, I estimate an OLS regression

Figure 1.14: Average Daily Consumption by Household Income



Note: Using 2004–2017 Nielsen Consumer Panel data, this figure plots the income bin coefficients from Equation (1.19), which regresses average daily household toilet paper consumption on household characteristics. Average daily consumption is computed by dividing total quantity purchased in a year by the number of days a household was active in the panel.

of consumption on household characteristics:

$$Y_i = \alpha + \beta X_i + \epsilon_i, \tag{1.19}$$

where Y_i is household i's average daily consumption and X_i is a vector of household characteristics. Figure 1.14 plots the income coefficients of an OLS regression including only income covariates and the coefficients when household characteristics are included. The graph illustrates that after controlling for covariates that plausibly cause increased consumption, income is not significantly correlated with consumption.

The above specification omits many other possible covariates that could be correlated with average daily consumption. When there are many possible variables that can be included, there is a risk of over-fitting. Elastic net regularization is a machine learning method that penalizes over-fitting and selects only the most predictive variables.

The elastic net solves the following minimization problem:

$$\min_{\beta} \|y - X\beta\|^2 + \lambda \left(\alpha\|\beta\|_1 + (1-\alpha)\|\beta\|_2^2\right), \tag{1.20}$$

where $\|\cdot\|_1$ is the L1 norm and $\|\cdot\|_2$ is the L2 norm. The OLS estimate is the β that solves the minimization problem with only the first term. The second term and third term provide penalties to shrink and select for the most predictive variables.

I set the mixing parameter α to be 0.5. When covariates are correlated in groups, lasso regression ($\alpha = 1$) tends to only select one and discard all other members of the group while ridge regression ($\alpha = 0$) tends to shrink correlated coefficients towards each other (Zou and Hastie 2005). Because some of the possible covariates form natural groups (e.g., all income bins or all markets), I chose $\alpha = 0.5$ since this tends to include or exclude groups together.

I estimate a 100-fold cross-validated elastic net regression to select the most predictive covariates. The resulting estimates selects many household characteristics including household composition, age, marital status, and race, but excludes almost all income and geographic coefficients.[48]

Appendix Tables

Table 1.22: Nielsen Consumer Panel Summary Statistics by Unit Price Regulation

	Without Regs		With Regs	
Variable	Mean	SD	Mean	SD
Household income ($000s)	55.65	30.75	59.60	31.70
Household size	2.53	1.43	2.61	1.49
Age	52.34	14.37	53.02	14.43
College Educated	0.37	0.48	0.41	0.49
Child present	0.33	0.47	0.32	0.47
Married	0.52	0.50	0.49	0.50
N (Household-Years)	488,461		246,263	

Note: Unweighted means and standard deviations are reported.

[48] Elastic net results are available upon request.

Table 1.23: Nielsen Consumer Panel Summary Statistics for Households that Move

	Reg to Reg		Reg to No Reg	
Variable	Mean	SD	Mean	SD
Household income ($000s)	60.03	29.14	62.32	28.60
Household size	2.17	1.20	2.24	1.20
Child present	0.18	0.39	0.17	0.38
Married	0.54	0.50	0.66	0.47
College Educated	0.56	0.50	0.60	0.49
Age	56.24	13.10	58.09	12.53
N (Household-Year)	26,393		6,027	

	No Reg to Reg		No Reg to No Reg	
Variable	Mean	SD	Mean	SD
Household income ($000s)	59.75	29.02	57.11	28.77
Household size	2.17	1.14	2.25	1.26
Child present	0.15	0.35	0.21	0.41
Married	0.64	0.48	0.58	0.49
College Educated	0.52	0.50	0.55	0.50
Age	59.30	12.34	55.36	13.08
N (Household-Year)	4,260		56,059	

Note: Unweighted means and standard deviations are reported.

Table 1.24: Robustness Test: Warehouse Club Entry on Bulk Buying

	5 Mi	10 Mi	15 Mi	20 Mi
	(1)	(2)	(3)	(4)
Post-Entry	−0.005	−0.007	0.001	0.005
	(0.006)	(0.007)	(0.008)	(0.008)
Post-Entry : 25-50k	0.012***	0.012**	0.019***	0.014**
	(0.005)	(0.006)	(0.006)	(0.005)
Post-Entry : 50-100k	0.014***	0.022***	0.025***	0.023***
	(0.004)	(0.005)	(0.005)	(0.006)
Post-Entry : >100k	0.017***	0.030***	0.035***	0.043***
	(0.006)	(0.008)	(0.009)	(0.010)
Household-ZIP FE's	Y	Y	Y	Y
Year-Qtr FE's	Y	Y	Y	Y
Demographic Controls	Y	Y	Y	Y
Observations	2,400,344	2,401,665	2,401,038	2,400,924
Adjusted R^2	0.428	0.428	0.428	0.428
Note:	*p<0.1; **p<0.05; ***p<0.01			

Note: This table uses 2004–2015 Nielsen Consumer Panel data at the household-quarter level. Coefficients are reported for Equation (1.9) which regresses households' quarterly bulk purchase shares on an indicator for warehouse club entry, an indicator for whether the household shops at a warehouse club, and an interaction term as well as household characteristics. Different distance cutoffs defining an "entry" are used for each regression. Household-ZIP code and year-quarter fixed effects are included. Projection weights are not used.

Chapter 2

A WORLD WITHOUT BORDERS REVISITED: THE IMPACT OF ONLINE SALES TAX COLLECTION ON SHOPPING AND SEARCH

By E. Mallick Hossain

2.1 Introduction

Online shopping has grown dramatically since 2000, reaching 11.4% of total retail sales in Q4 of 2019 (U.S. Census Bureau 2020). Reasons for this growth include lower costs (travel, time, etc.), higher convenience, and more variety (Thau 2013; Arnott 2016). Until 2018, traditional brick-and-mortar stores were at a structural disadvantage because they had to collect sales taxes while their online competitors did not. As a result, online retailers could offer consumers, on average, a 7% discount compared to brick-and-mortar retailers. This discount came strictly at the expense of state and local tax revenues, with losses ranging from \$8 to \$33 billion in 2018 (U.S. Supreme Court 2018).

In this paper, I estimate consumers' price elasticity using price increases generated by sales taxes on online purchases. Due to pressure from state and local governments and aggressive fulfillment center expansion, online retailers (particularly Amazon) collect sales taxes in many states. Combining data on online shopping with local tax rates and Amazon's tax collection behavior across states over time, I estimate how consumers' online purchasing changes when sales taxes are collected online. Furthermore, I extend this analysis to include measures of online browsing as well as overall household spending (including offline expenditures) to estimate whether consumers' search behavior or their composition of online and offline spending changes in reponse to online sales tax collection.

I combine online shopping data with offline sales tax rates to estimate whether areas with higher tax rates respond more strongly to Amazon's sales tax collection. Previous research uses a variety of approaches to determine how sensitive consumers are to tax rates. Table 2.1 summarizes elasticity estimates from previous research. In the offline environment, research leverages cross-border variation in tax rates and estimates a wide range of elasticities from -30 to -0.2 (Asplund, Friberg, and Wilander 2007; Agarwal, Marwell, and McGranahan 2017; Davis 2011; Agarwal et al. 2017; Mikesell 1970). In the online environment, elasticity estimates range from -6 to 0, but early efforts often used data from before 2001, i.e., before the mass adoption of the internet and before groundbreaking innovations like rating systems and free shipping were effectively implemented (Scanlan 2007; Ballard and Lee 2007; Alm and Melnik 2005; Goolsbee 2000). Recent work leverages detailed online shopping data, but often is limited to particular websites, product categories, or states (Einav et al. 2014; Anderson et al. 2010; Ellison and Ellison 2009; Hu and Tang 2014). My paper extends the work of Baugh, Ben-David, and Park (2018) and Houde, Newberry, and Seim (2017) by incorporating data on browsing and total household expenditures to present a fuller picture of how household behavior changes in response to Amazon's sales tax collection.

I use a differences-in-differences approach to estimate a household's price elasticity. The expansion of Amazon's warehouse network and the passage of state laws requiring online sales tax collection generate variation in Amazon's tax liability across states over time. As a result, I am able to examine how household behavior changes after Amazon begins collecting sales taxes.

I find that consumers reduce their Amazon spending by about 1.9% for each percentage point of sales tax applied online. Given an average sales tax rate of 7%,

Table 2.1: Estimated Elasticities

Elasticity Type	Paper	Estimate
Cross-border Price-Expenditure	Asplund et al. (2007) – Foreign price	0.2 to 0.5
	Asplund et al. (2007) – Domestic price	-0.2 to -1.3
	Agarwal et al. (2017)	-2 to -30
	Davis (2011)	-2.2 to -3.6
	Agarwal et al (2017)	-2.3
	Mikesell (1970)	-6.3
Tax-Purchase	Scanlan (2007)	0.0
	Ballard and Lee (2007)	-0.2
	Alm and Melnik (2005)	-0.5
	Einav et al. (2014)	-1.8
	Goolsbee (2000)	-2.3
Tax-Quantity	Anderson et al. (2010)	-1.9 to -2.9
	Ellison & Ellison (2009)	-6
Tax-Expenditure	Baugh et al. (2018)	-1.2 to -1.4
	Houde et al. (2017)	-1.3
	Hu and Tang (2014)	-3.75 to -4.5

Note: Early research focused on how taxes influenced the binary decision of whether or not to make an online purchase. Subsequent research has looked at how sales taxes affect actual online expenditures or quantities purchased. In order to delineate between these, I use "tax-purchase elasticity" to refer to the effect on the purchase decision while "tax-quantity elasticity" and "tax-expenditure elasticity" refer to the effect on online purchase quantities and expenditures, respectively.

this elasticity translates into a 13% decline in retail spending on Amazon. Consumers also increase their spending on Amazon's taxed competitors by 1% per percentage point of sales tax collected by Amazon. This is one of the first papers to explicitly incorporate how consumer search behavior changes in response to tax changes. Even though consumers do shift their spending from Amazon to its competitors, I find no evidence that consumers' browsing habits are significantly affected by Amazon's sales tax collection. This is also one of the first papers to examine whether households shift their spending offline in response to online sales tax collection. I find no evidence that consumers make such a shift.

This paper is organized as follows. Section 2.2 describes the data. Section 2.3 analyzes how online spending responds to Amazon sales tax collection. Section 2.4 examines whether online browsing activity is affected by Amazon sales tax collection.

Section 2.5 analyzes tax responsiveness across all consumer spending modes and Section 2.6 concludes.

2.2 Data Description

In this section, I describe the data used for my analysis and give a brief overview of their respective features.[49] comScore's Web Behavior database provides information on household online shopping and browsing behavior. Nielsen's Consumer Panel data provides information on household shopping and purchasing decisions. Finally, Tax Data Systems provides information on local sales tax rates.

comScore Web Behavior Database

I primarily use the comScore Web Behavior database, which contains the online browsing and transaction activity of households that opt-in to have their internet activity collected by comScore. The browsing data records how many minutes were spent and how many pages were viewed on each website. The transaction data records the website, product name, product category, price, quantity, and basket total (including shipping and taxes) of the purchase. The comScore data capture all online activity of a household and is not limited to particular goods or retailers, in contrast with previous research. I use this breadth of information to capture whether households substitute to other retailers and estimate how their aggregate online spending changes when sales taxes are collected online.

I restrict my sample to households that have complete demographic information and remove any purchases in categories which Amazon is a not competitor (e.g., no plane tickets, dating services, etc.) and focus on products that cost between $1 and $500 (in nominal dollars).[50] These filters reduce the original sample of about

[49] Researcher's own analyses derived based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

[50] This restriction is primarily to screen out extreme prices that may be generated by the comScore monitoring software. For example, a $20 item reduced to $15 may mistakenly be recorded as $2015 because of how the price information is captured.

576,000 households down to about 206,000 households and a total of about 2 million transactions.

Table 2.2: comScore Panel Summary Statistics

Statistic	Mean	Median	St. Dev.	Min	Max	Census (2016)
Household Size	3.04	3	1.40	1	6	2.53
Age	46.65	47	12.59	19	65	51.9
Income	59.01	62.50	31.66	7.50	100.00	59.04
Child Present	0.64	1	0.48	0	1	0.42
Hispanic	0.13	0	0.34	0	1	0.13
Sales Tax	0.07	0.07	0.02	0.00	0.11	-
N (Household-Years)			261,416			

Note: Age and income are reported in bins, so the midpoint of each bin is used. "Child Present" indicates whether a child is present in the household. Census data comes from "Historical Households Tables" and "Income and Poverty in the United States: 2016."

Table 2.2 shows the summary statistics of the panelists. The average household size is 3, average age is 47, and average income is \$59,000. About 64% of the households have a child and 13% are Hispanic. Households face an average sales tax rate of 7%. Overall, panelists are similar to the general population, but are slightly larger and more likely to have kids. Households rotate through comScore's panel with a median tenure of 12 months (the 25th percentile is 9 months and the 75th percentile is 12 months).

Table 2.3 shows the summary statistics for all 2 million transactions. The average real price of products (in December 2016 dollars) is \$39. 28% of products are purchased on Amazon, 42% from the website of a brick-and-mortar store and the remaining products are purchased from another online retailer.

Table 2.4 reports the summary statistics of browsing on shopping websites. The average household spends about two and a half hours per month on shopping websites. Browsing on Amazon's untaxed competitors accounts for about one and a half hours, browsing on Amazon accounts for about 20 minutes and the remaining 40 minutes is on Amazon's taxed competitors. There is a wide range of shopping behavior with many households spending no time on shopping websites in a given month and others spending up to five hours per day.

Table 2.3: comScore Transaction Summary Statistics

Statistic	Mean	Median	St. Dev.	Min	Max
Real Product Price	39.09	20.82	57.06	1.00	608.90
Sales Tax	0.07	0.07	0.02	0.00	0.11
Amazon Purchase	0.28	0	0.45	0	1
Offline Amazon Competitor	0.42	0	0.49	0	1
Online Amazon Competitor	0.30	0	0.46	0	1
N	2,001,485				

Note: Prices are deflated to December 2016 price levels using the CPI. "Sales Tax" indicates average local sales tax rate, not the average sales tax paid on online transactions.

Table 2.4: comScore Browsing Summary Statistics (Minutes)

Statistic	Mean	Median	St. Dev.	Min	Max
Total	152.50	51.7	388.33	0	10,196
Amazon	18.52	0	83.19	0	8,864
Untaxed Competitor	93.46	16	346.00	0	10,176
Taxed Competitor	40.51	8	100.53	0	9,810
N (Household-Months)		2,559,012			

Note: Using 2006–2016 comScore Web Behavior data, this table reports the distribution of monthly browsing durations, in minutes, on shopping websites.

Nielsen Consumer Panel Data

I use the Nielsen Consumer Panel data from 2004–2016. This is a panel of about 178,000 unique households. I observe about 40,000 households each year from 2004–2006 and about 60,000 households each year from 2007–2016. Households scan all items that they purchase and then input information about quantities, prices, date of purchase, and store. Nielsen retains about 80% of its panel from year to year with the mean and median tenure of a household being four and three years, respectively.

Nielsen computes projection weights to ensure their sample is nationally representative. Weights are calculated to match population moments based on household size, income, age, race, ethnicity, education, occupation, and presence of children. All analyses use these projection weights unless otherwise stated. Table 2.5 presents descriptive statistics for households in the sample.

Table 2.5: Nielsen Consumer Panel Summary Statistics

Variable	Mean	SD	25th Pctile	75th Pctile	Census
Household income ($000s)	56.53	31.41	27.5	85	59.04
Household size	2.55	1.45	1	3	2.53
Age	52.62	14.34	41.5	63	51.9
College Educated	0.38	0.48	0	1	0.37
Child present	0.33	0.47	0	1	0.42
Married	0.50	0.50	0	1	0.48
N (Household-Years)			637,493		
N (Households)			154,352		

Note: Data are weighted for national representativeness.

Unlike the comScore data, store identities are anonymized, but they are classified into broad categories such as "Grocery Stores," "Electronics Store," and "Online Shopping." This categorization is enough to conduct a similar analysis as I do with the comScore data, with the caveat that the Nielsen data primarily focuses on basic household goods, so items like electronics and apparel will not be captured. For retailers with both online and offline presence, Nielsen classifies them separately. For example, if Firm X has both offline stores and a website, detergent purchased from a Company X store will be from a different "retailer" than detergent purchased from CompanyX.com. In the first case, the retailer would be classified as a "Discount Store" and in the second, the retailer is classified as "Online Shopping."

Because of their addictive properties and laws regulating purchasing these products online, I exclude tobacco and alcohol products. I remove households with a student or military head of household as well as those with an annual income of less than $5,000. Only about 2% of households are excluded and I use the remaining 154,000 households for my analysis. See Appendix 2.A for further details of sample construction.

Additional Data Sources

I obtain state, county, and local sales tax rates from Tax Data Systems, now part of Thomson Reuters. These data contain monthly tax rates at the ZIP code level. I compile information on state law changes and agreements with Amazon under which states began collecting taxes for online transactions. This information was gathered from a wide range of local, state, and national news sources. Prior to 2018, most states did not require online retailers to collect sales taxes. They have only been able to collect sales taxes from Amazon because of separate agreements or because Amazon has opened warehouses in their state.[51] Before 2006, Amazon only collected sales taxes in Kansas, Kentucky, North Dakota, and Washington. By the end of 2016, Amazon collected sales taxes in an additional 25 states: Alabama, Arizona, California, Colorado, Connecticut, Florida, Georgia, Illinois, Indiana, Maryland, Massachusetts, Michigan, Minnesota, Nevada, New Jersey, New York, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Texas, Virginia, West Virginia, and Wisconsin as well.

Figure 2.1 shows the wide variation in sales tax rates across the United States in December 2016. About 80% of counties have combined state and local sales tax rates between 5.3% and 8.3%. However, sales tax rates range from 0% to over 10% in Louisiana.

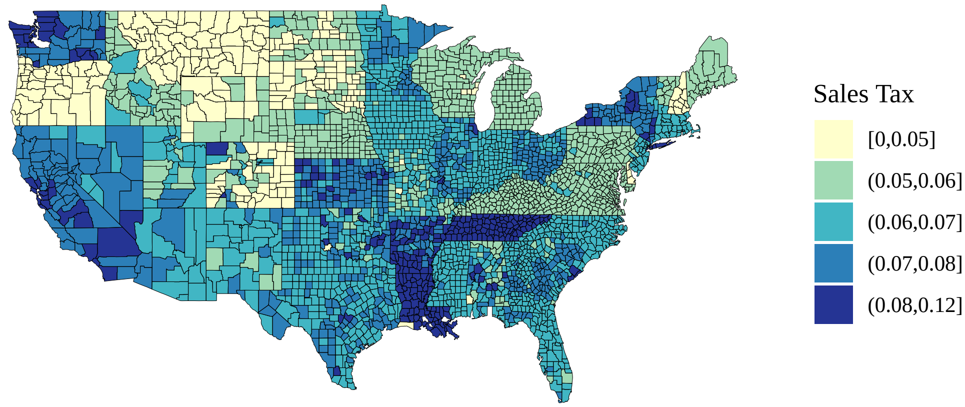
2.3 Amazon Sales Tax Collection and Online Spending

When shopping online, households can purchase from Amazon or one of its competitors. Amazon has two types of competitors: taxed and untaxed. Amazon's taxed competitors consist of traditional brick-and-mortar retailers, like Walmart and Target, who collect sales tax since they have physical locations across the country. Amazon's untaxed competitors are other online retailers, like Overstock.com or Etsy.com, which do not have physical locations across the country (generally just a headquarters location).

When Amazon begins collecting sales tax, consumers could respond in a variety of ways. First, they may not change their behavior and purchase on Amazon like usual (maybe not even notice the sales tax). Second, they could switch to one of Amazon's competitors. They could switch to an untaxed competitor if they value the tax savings

[51] *National Bellas Hess v. Illinois* (1967) and *Quill Corp. v. North Dakota* (1992) ruled that retailers did not have to collect sales taxes in states where they did not have a physical presence. The Court held that tabulating tax liabilities for over 6,000 different tax jurisdictions would place an undue burden on many of these firms (Atkins 2005). The Supreme Court overturned these cases in *South Dakota v. Wayfair Inc.* (2018). Before 2018, consumers were supposed to self-report any unpaid taxes to the tax authorities, but compliance and enforcement were low, so these transactions were effectively tax-free (Manzi 2015).

Figure 2.1: Sales Tax Rates by County (Dec 2016)



Note: Using December 2016 Tax Data Systems data, this figure plots the spatial distribution of sales tax rates.

or possibly a taxed competitor that offers a better selection or lower tax-inclusive prices. I examine each of these in turn to see how spending on Amazon changes and how spending at Amazon's taxed and untaxed competitors changes after Amazon begins collecting sales taxes.

I use a differences-in-differences specification to identify the effect of Amazon's sales tax collection on a household's online spending:

$$Y_{ht} = \alpha + \beta AmazonCollect_{ht} * \tau_{ht} + \lambda_h + \lambda_t + \epsilon_{ht}, \tag{2.1}$$

where Y_{ht} measures real expenditures of household h in month t. $AmazonCollect_{ht}$ indicates whether Amazon collects sales tax (determined by month and state of residence of the household). τ_{ht} is the local sales tax rate.[52] Household and time fixed effects are captured by λ_h and λ_t. β is the coefficient of interest measuring how

[52]The local tax rate could also be included separately, but given the household fixed effect, this would only be identified off of changes in local tax rates which are relatively infrequent and when they do happen, are small.

spending changes after Amazon begins collecting sales taxes relative to what would have been expected had they not started collecting sales taxes. All standard errors are clustered at the state level because once Amazon collects sales taxes, it collects them across the whole state.

The policy of whether or not Amazon collects sales tax in a particular state or county is plausibly exogenous to the household spending decision. Often, it is prompted by the opening of an Amazon warehouse in the state, but in a few cases, it is because of a change in state law. While there is a chance that these changes could be related to underlying economic fundamentals, Baugh, Ben-David, and Park (2018) show that sales tax collection by Amazon is not significantly related to state GDP growth, household income changes, or consumption declines. About 10% of households experience a change in Amazon's sales tax collection while they are in the sample.

Table 2.6 reports the estimation results. Columns (1) – (3) only include an indicator for whether or not Amazon collects sales tax and then columns (4) – (6) allow for the response to vary with the sales tax rate. Column (1) demonstrates that Amazon spending decreases by an average of \$0.422 after Amazon collects sales taxes. Given that average monthly spending on Amazon is \$3.30 and the average sales tax rate is 6.8%, this equates to an elasticity of $\frac{-0.422/3.30}{0.068} = -1.88$. Columns (2) and (3) show that spending on Amazon's untaxed competitors does not significantly change while spending on its taxed competitors increases by \$0.549, implying a cross-elasticity of 1.14. Column (4) shows that the spending decreases on Amazon are stronger in areas with higher sales tax rates and the implied elasticity is a similar -1.87.[53] As before, there is no significant response on Amazon's untaxed competitors and a marginally significant increase on Amazon's taxed competitors. The spending increase on Amazon's taxed competitors implies an elasticity of 0.97. Overall, after Amazon collects sales taxes, households reduce their Amazon spending and increase their spending on Amazon's taxed competitors.

My estimated elasticity of -1.88 is higher (in magnitude) than similar estimates from Baugh, Ben-David, and Park (2018) and Houde, Newberry, and Seim (2017). My estimate differs from Houde, Newberry, and Seim (2017) for two reasons. First,

[53] 6.173 / 3.30 = -1.87

Table 2.6: Online Spending Response to Amazon Sales Tax Collection

	Amazon	Untaxed Sites	Taxed Sites	Amazon	Untaxed Sites	Taxed Sites
	(1)	(2)	(3)	(4)	(5)	(6)
Collect	-0.422^{**}	-0.112	0.549^{**}			
	(0.183)	(0.150)	(0.249)			
Collect * Tax Rate				-6.173^{**}	-2.043	6.811^{*}
				(2.715)	(2.165)	(3.545)
Household FE	Y	Y	Y	Y	Y	Y
Month-Year FE	Y	Y	Y	Y	Y	Y
Mean Spending	3.30	4.86	7.01	3.30	4.86	7.01
Mean Tax	0.068	0.068	0.068	0.068	0.068	0.068
Observations	5,076,040	5,076,040	5,076,040	5,076,040	5,076,040	5,076,040
Adjusted R^2	0.124	0.203	0.184	0.124	0.203	0.184

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Note: Using 2006–2016 comScore data on household-month online expenditures, this table reports the estimation results of Equation 2.1. “Collect” is a dummy variable indicating whether Amazon collected sales tax in a particular household-month. “Tax Rate” measures the local sales tax rate faced by a household in a particular month. All expenditures are real expenditures, deflated to December 2016 using the CPI. “Taxed Sites” refers to websites of retailers that have offline stores. “Untaxed Sites” are online-only retailers with no offline stores. Standard errors are clustered at the state level.

while I use the same underlying data, I extend my sample for another three years through 2016, which doubles the number of states in which Amazon begins collecting sales tax from 12 to 25. Second, their analysis aggregates the data to the county-year level, while I aggregate the data to the household-month level. I get nearly identical estimates if I aggregate to the county-year level and limit my sample to 2006–2013. My estimate also differs from Baugh, Ben-David, and Park (2018) likely because of differences in the underlying data. First, my analysis spans 2006–2016 compared to 2011–2015, which adds an additional six states to my analysis.[54] Restricting my analysis to 2011–2015 generates a slightly smaller elasticity of −1.66, but this is still

[54] Baugh, Ben-David, and Park (2018) also restrict their analysis to households that spend more than $200 on Amazon in 2011, but their Appendix B shows that removing this filter does not impact their estimate.

higher than the -1.2 to -1.4 estimated in Baugh, Ben-David, and Park (2018). The other possible contributor is the composition of our samples. The comScore data captures all online activity on a household's computer and panelists are recruited to provide a representative measure of US internet users' activity. On the other hand, the data used in Baugh, Ben-David, and Park (2018) is from an online account aggregator that likely targets younger, tech-savvy users interested in managing their finances effectively.[55] These users are probably more likely to shop online and may be less likely to switch away from Amazon. This assertion is supported by comparing the average monthly spending on Amazon between the two samples. The average monthly Amazon spending of a comScore user is only \$3.30, but this increases to \$12.20 when restricting to only households that have made purchases on Amazon. In comparison, the average household in Baugh, Ben-David, and Park (2018) spends \$39. Overall, my estimate is higher than previous estimates because I incorporate more recent data and (arguably) a more representative sample of online shoppers.

2.4 Amazon Sales Tax Collection and Online Browsing

The previous section shows that consumers are spending less on Amazon and more on Amazon's taxed competitors. Do these changes in spending translate into changes in search behavior? I estimate Equation 2.1 with Y being minutes spent on Amazon or one of its competitors' websites.

Table 2.7 shows the results of this estimation. Our previous results indicate that households reduce their spending on Amazon only after sales tax is collected online. Because of this, we might expect that this reduced shopping activity would translate into reduced overall activity, measured in time spent on the website. Overall, I find no evidence that search on Amazon or its competitor websites is significantly affected by Amazon collecting sales tax. The lack of a significant browsing response may indicate that consumers are not changing their search behavior, but are simply switching their

[55] One of the most popular financial aggregators, Mint.com, is reported to have a primarily young, male demographic. 71% of users were male and 64% were under 30 years of age back in 2008 (Perez 2008).

Table 2.7: Online Browsing Response to Amazon Sales Tax Collection

	Minutes Browsed					
	Amazon	Untaxed Sites	Taxed Sites	Amazon	Untaxed Sites	Taxed Sites
	(1)	(2)	(3)	(4)	(5)	(6)
Collect	0.076	0.387	0.512			
	(0.435)	(0.977)	(0.640)			
Collect * Tax Rate				0.369	6.317	7.160
				(5.695)	(12.840)	(9.243)
Household FE	Y	Y	Y	Y	Y	Y
Month-Year FE	Y	Y	Y	Y	Y	Y
Mean Browsing	11.71	66.66	27.98	11.71	66.66	27.98
Mean Tax	0.068	0.068	0.068	0.068	0.068	0.068
Observations	5,076,040	5,076,040	5,076,040	5,076,040	5,076,040	5,076,040
Adjusted R^2	0.463	0.410	0.386	0.463	0.410	0.386
Note:	*p<0.1; **p<0.05; ***p<0.01					

Note: Using 2006–2016 comScore data on household-month online browsing, this table reports the estimation results of Equation 2.1. "Collect" is a dummy variable indicating whether Amazon collected sales tax in a particular household-month. "Tax Rate" measures the local sales tax rate faced by a household in a particular month. All browsing is in minutes. Standard errors are clustered at the state level.

purchases away from Amazon since it no longer has a sales tax advantage.

Households' relative unresponsiveness in search effort contrasts with the findings of Einav et al. (2014), which finds that when buyers realize sales taxes are added, they back out of the transaction. This could be due to differences in user search between Amazon and eBay. On Amazon, the products are listed at a fixed price while on eBay, a share of items are sold at auction. In 2010 (the data used in Einav et al. (2014)), 40% of eBay sales were from auctions (eBay.com 2011). Given the risk of losing the auction, customers may be more likely to search on eBay relative to Amazon, where there is no risk of losing the purchase. Even if there is an effect that I cannot detect, it is likely to be small changes in browsing time, which could be generated by the extra effort needed to complete the purchase (e.g., the time needed to enter in address and credit card information).

Overall, households reduce their pre-tax expenditures on Amazon and shift to

Amazon's taxed competitors. In the next section, I examine whether households change their overall expenditures with a focus on whether their offline expenditures change in response to Amazon collecting sales tax.

2.5 Total Consumer Expenditure Analysis

The comScore data suggest that consumers reduce their Amazon expenditures when Amazon begins collecting sales tax, but they do not spend or browse significantly more on Amazon's online competitors. The comScore analysis is limited to examining only online transactions and activity. Using Nielsen's Consumer Panel Data, this section examines how households' overall spending changes in response to Amazon's sales tax collection. Households scan all of the items that they purchase for at-home consumption and input information about their shopping trip, including whether the purchase was from an online or offline retailer.[56] Using these data, I can determine, for the common household items that are tracked, whether Amazon's sales tax collection changes overall household expenditures and whether households are shifting spending offline in response.

To identify changes in consumer expenditures, I estimate Equation 2.1 where Y is the expenditures in either the online-only (likely untaxed) channel or offline channel. Since Amazon is most competitive in delivering less perishable, non-food products, I also separately examine whether online or offline non-food spending is affected. About 45% of households experience a change in Amazon's sales tax collection during their tenure.

Table 2.8 shows that a range of household spending groups are unaffected by Amazon's sales tax collection. Columns 1–3 show that there are no significant changes in total, online, or offline spending after Amazon begins collecting sales taxes. Even if there is an effect that I am not detecting, it is likely quite small. Part of this result is due to the fact that online shopping has low penetration into grocery and household non-durables (as indicated by online monthly spending averaging $4). Even when analyzing only non-food items, where online shopping is most likely, there is no

[56] Retailers are anonymized in Nielsen, so I cannot identify whether an online purchase is made at a taxed or untaxed website as I could in the comScore data.

Table 2.8: Household Spending Response to Amazon Sales Tax Collection

	Real Spending				
	Total	Online	Offline	Online Non-Groc	Offline Non-Groc
	(1)	(2)	(3)	(4)	(5)
Collect	−0.682	0.036	−0.719	0.001	0.118
	(1.734)	(0.169)	(1.772)	(0.095)	(0.458)
Household FE	Y	Y	Y	Y	Y
Month-Year FE	Y	Y	Y	Y	Y
Mean Spending	316.39	4.7	311.69	2.93	94.74
Mean Tax	0.068	0.068	0.068	0.068	0.068
Observations	7,792,355	7,792,355	7,792,355	7,792,355	7,792,355
Adjusted R^2	0.531	0.331	0.532	0.226	0.399

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Note: Using 2006–2016 Nielsen Consumer Panel data, this table reports the estimates from regressing monthly spending on an indicator for Amazon sales tax collection ("Collect") as well as household and month-year fixed effects and household demographics. Standard errors are clustered at the state level.

significant effect and the range of possible changes is quite small.

Because online shopping for household non-durables is relatively infrequent, the numerous months with no online expenditures may be masking changes in online purchases when they are made. To check this, I estimate the same regression, but only using months when purchases are made (i.e., conditional on making a purchase).

Table 2.9 reports the estimates conditional on making a purchase. Even after conditioning on making a purchase, online spending does not significantly change after Amazon collects sales taxes. Overall, there is no evidence of households shifting their spending offline in response to Amazon's tax collection.

Table 2.9: Household Spending Response to Amazon Sales Tax Collection (Positive Purchases Only)

	Real Spending				
	Total	Online	Offline	Online Non-Groc	Offline Non-Groc
	(1)	(2)	(3)	(4)	(5)
Collect	−0.682	1.063	−0.699	−0.107	0.128
	(1.734)	(1.106)	(1.755)	(0.974)	(0.458)
Household FE	Y	Y	Y	Y	Y
Month-Year FE	Y	Y	Y	Y	Y
Mean Spending	316.39	72.57	312.2	52.94	99.55
Mean Tax	0.068	0.068	0.068	0.068	0.068
Observations	7,792,355	482,702	7,782,003	408,811	7,464,450
Adjusted R^2	0.531	0.377	0.532	0.234	0.392

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Note: Using 2006–2016 Nielsen Consumer Panel data, this table reports the estimates from regressing monthly spending (conditional on have positive spending that month) on an indicator for Amazon sales tax collection ("Collect") as well as household and month-year fixed effects and household demographics. Standard errors are clustered at the state level.

2.6 Future Research and Conclusion

Using data covering a broad range of online shopping activity, I find that consumers reduce their pre-tax spending on Amazon by about 1.9% for every percentage point of sales tax that Amazon collects. Furthermore, I find that households increase their spending on Amazon's taxed competitors by 1% for each percentage point of sales tax Amazon collects. Even though households change their spending, they do not significantly change their search behavior on Amazon or its competitors. Finally, I find no evidence that households shift any of their spending offline after Amazon collects sales tax.

In light of the recent Supreme Court case, *South Dakota v. Wayfair* which increases state enforcement of sales tax collection online, state and local governments can expect a revenue boost because consumers are unlikely to shift their spending to untaxed channels. However, local policymakers and businesses will need to find other approaches if they want to encourage shoppers to move back offline. Online shopping is here to stay and more empirical work will be necessary to understand how traditional offline retailers can adapt to increased online shopping.

References

Agarwal, Sumit, Nathan Marwell, and Leslie McGranahan (2017). "Consumption responses to temporary tax incentives: Evidence from state sales tax holidays". In: *American Economic Journal: Economic Policy* 9.4, pp. 1–27.

Agarwal, Sumit et al. (2017). "Tax Differential and Cross-Border Shopping: Evidence from Singapore". In: *Georgetown McDonough School of Business Research Paper No. 3038550*.

Alm, James and Mikhail I Melnik (2005). "Sales taxes and the decision to purchase online". In: *Public Finance Review* 33.2, pp. 184–212.

Anderson, Eric T et al. (2010). "How sales taxes affect customer and firm behavior: The role of search on the Internet". In: *Journal of Marketing Research* 47.2, pp. 229–239.

Arnott, Grant (May 11, 2016). *Why People Buy Online vs In Store – Not What You Think*. URL: https://www.powerretail.com.au/news/online-vs-store-research/.

Asplund, Marcus, Richard Friberg, and Fredrik Wilander (2007). "Demand and distance: evidence on cross-border shopping". In: *Journal of public Economics* 91.1-2, pp. 141–157.

Atkins, Chris (July 2005). *Important Tax Cases: Quill Corp. V. North Dakota and the Physical Presence Rule for Sales Tax Collection*. English. Tax Foundation. URL: https://taxfoundation.org/important-tax-cases-quill-corp-v-north-dakota-and-physical-presence-rule-sales-tax-collection.

Ballard, Charles L and Jaimin Lee (2007). "Internet purchases, cross-border shopping, and sales taxes". In: *National Tax Journal*, pp. 711–725.

Baugh, Brian, Itzhak Ben-David, and Hoonsuk Park (2018). "Can Taxes Shape an Industry? Evidence from the Implementation of the "Amazon Tax"". In: *The Journal of Finance* 73.4, pp. 1819–1855.

Davis, Lucas W (2011). “The effects of preferential VAT rates near international borders: evidence from Mexico”. In: *National Tax Journal* 64.1, p. 85.

eBay.com (2011). *2010 Annual Report*. URL: https://d18rn0p25nwr6d.cloudfront.net/CIK-0001065088/f62994ae-7cd0-4396-a4b4-450b21fa8502.pdf.

Einav, Liran et al. (2014). “Sales taxes and internet commerce”. In: *American Economic Review* 104.1, pp. 1–26.

Ellison, Glenn and Sara Fisher Ellison (2009). “Tax sensitivity and home state preferences in internet purchasing”. In: *American Economic Journal: Economic Policy* 1.2, pp. 53–71.

Goolsbee, Austan (2000). “In a world without borders: The impact of taxes on Internet commerce”. In: *Quarterly Journal of Economics* 115.2, pp. 561–576.

Houde, Jean-François, Peter Newberry, and Katja Seim (2017). *Economies of Density in E-Commerce: A Study of Amazon’s Fulfillment Center Network*. Working Paper 23361. National Bureau of Economic Research.

Hu, Yu Jeffrey and Zhulei Tang (2014). “The impact of sales tax on internet and catalog sales: Evidence from a natural experiment”. In: *International Journal of Industrial Organization* 32, pp. 84–90.

Manzi, Nina (Apr. 2015). *Use Tax Collection on Income Tax Returns in Other States*. Research rep. 600 State Office Building St. Paul, MN 55155: Minnesota House of Representatives. URL: http://www.house.leg.state.mn.us/hrd/pubs/usetax.pdf.

Mikesell, John L (1970). “Central cities and sales tax rate differentials: The border city problem”. In: *National Tax Journal* 23.2, pp. 206–213.

Perez, Sarah (Oct. 2008). *Who Really Uses Mint.com (And Other Banking 2.0) Services Anyway?* English. readwrite. URL: https://readwrite.com/2008/10/09/who_really_uses_mint_anyway/.

Scanlan, Mark A (2007). “Tax sensitivity in electronic commerce”. In: *Fiscal Studies* 28.4, pp. 417–436.

Thau, Barbara (Oct. 8, 2013). *New Study Reveals Why Consumers Really Shop Online (Surprise: It Isn’t Low Prices)*. URL: https://www.forbes.com/sites/barbarathau/2013/10/08/why-consumers-really-shop-online/.

U.S. Census Bureau (2020). *Quarterly Retail E-Commerce Sales*. English. URL: https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf.

U.S. Supreme Court (2018). "South Dakota v. Wayfair Inc." In: *U.S.* 585.

2.A Data Appendix

comScore Web Behavior

This section provides more details about how I prepared the comScore data for analysis. All comScore data was obtained directly from Wharton Research Data Services (WRDS). I drop any households with incomplete demographic information. Additionally, I remove any households whose ZIP codes do not match with the Census Bureau's 2010 Zip-to-County Relationship file. Table 2.10 shows that 35% of households remain based on these filters. The low retention rate is primarily a majority of households in the comScore data just browse the internet and make no online purchases while they are in the sample.

Table 2.10: comScore Sample

Step	HH
Starting HH:	586, 420
Complete demographics:	585, 867
Valid ZIPs:	576, 457
Made Online Purchase:	206, 435

Note: Table reports the number of households remaining in sample after each step of data cleaning.

For online transactions, I remove any transactions that are recorded for the same visit, same price, same time, and same product as a duplicate record. I also restrict my sample to products in categories that Amazon competes in, which excludes travel, dating, and financial products. Furthermore, I drop any transactions where the price is missing, less than $1 or greater than $500. Then, I remove any products that were sold by websites that do not feasibly compete with Amazon (e.g. daysinn.com or date.com).[57] Table 2.11 shows that most transactions are omitted because they

[57] The full list of domains is available in the replication code.

are duplicates or in non-Amazon competitive categories. The remaining portion of transactions are removed because the household did not make any Amazon purchase while they were in the sample. Overall, 41% of all transactions are made in Amazon competitive categories.

Table 2.11: comScore Transactions Data Construction

Step	Transactions
Starting Transactions:	4,934,867
Unduplicated Transactions:	3,956,424
Amazon Categories:	2,478,115
Invalid Prices:	2,269,680
Invalid Domains:	2,021,800

Note: Table reports the number of transactions remaining after each step of data cleaning.

Nielsen Consumer Panel Data

This section provides more details about how I prepared the Nielsen Consumer Panel Data for analysis. I download the data directly from the University of Chicago Kilts Center for Marketing. I then remove any households with a military or student head of household or those that are making less than $5,000 annually. Table 2.12 shows that only 2% of households are removed by these criteria.

Table 2.12: Homescan Sample Construction

Step	HH
Starting HH:	158,004
Exclude military and students:	155,256
Exclude Households under 5k:	154,352

Note: Table reports the number of households remaining after each step of data cleaning.

In the purchases data, I remove any alcohol and tobacco products as well as product categories that Nielsen has not consistently tracked over the full 2006–2016 period (i.e., "deferred modules"). I also exclude products that do not have Universal Product Codes (i.e., "magnet modules"). Finally, I remove any products with a recorded price of zero.

One final note for those familiar with Nielsen's data products. Theoretically, I could use the Scanner data to determine whether an "online" retailer and another retailer share the same parent company (for the set of retailers that Nielsen track in their data). Unfortunately, none of the "Online Shopping" retailer codes are present in the Nielsen Scanner Data, so I cannot distinguish if the online retailer is the website of a traditional brick-and-mortar store (and thus taxed) or a stand-alone online store (and thus untaxed).

Chapter 3

MADE FROM SCRATCH: HOW SNAP BENEFITS AFFECT LOTTERY SALES

By E. Mallick Hossain and Jason Sockin

3.1 Introduction

State lotteries are an important source of revenue for state budgets, yet it is well documented that state lotteries are regressive (Price and Novak 2000; Combs, Kim, and Spry 2008) and appeal to low-income households (Haisley, Mostafa, and Loewenstein 2008; Blalock, Just, and Simon 2007). While lotteries are also a common form of entertainment (and possibly investment) for households, lotteries cannot substitute for food or other necessary goods and services. In order to assist low-income households in purchasing groceries, the U.S. government provides subsidies through the Supplemental Nutrition Assistance Program (SNAP). However, these subsidies do not translate into increased food consumption in a one-to-one fashion (Hastings and Shapiro 2018). Lottery purchases are often funded by reductions in non-gambling expenditures (Kearney 2005) and in Pennsylvania, 41% of lottery sellers are also stores where low-income households can purchase food using their SNAP benefits.

Our paper answers the question: Do SNAP benefits subsidize gambling? Because SNAP benefits effectively operate as cash for purchasing groceries, the cash that would have otherwise been spent on groceries can thus be used towards funding other activities, such as gambling. On the other hand, if they increase grocery spending, households may reduce their shopping trips and therefore gamble less frequently. To answer this question, we analyze the universe of state lottery tickets sold by retailers in Pennsylvania, and estimate how responsive lottery ticket sales are to the prevalence of SNAP-benefit households and the ability of stores to accept SNAP benefits. Previous research estimates that households spend 50–60% of their SNAP benefits on food (Hastings and Shapiro 2018)—or 40–50% on non-food expenditures—suggesting that a pass-through of SNAP benefits to gambling behavior is plausible and possibly non-trivial. Other research has found that other government transfers affect "sin" behaviors with unemployment insurance generating increases in drinking, but decreases in smoking (Lantis and Teahan 2018; Fu and Liu 2019).

Our paper is the first to estimate the relationship between SNAP benefits and lottery sales as well as present new stylized facts about the overlap between lottery retailers and SNAP-eligible retailers. Previous work has shown that the majority of SNAP benefits are spent on food (Hastings and Shapiro 2018), but little is known about what other products households spend SNAP on. Furthermore, lottery purchases are often funded by reductions in food and housing expenditures (Kearney 2005). For low-income households, the marginal benefits of these expenditures may be particularly high.

Second, we present a novel dataset on the universe of monthly store-level lottery ticket sales for Pennsylvania from October 2002 through March 2019. This data illustrates the rich diversity in the stores that sell lottery tickets, the kinds of lottery tickets that are most popular, and where these stores are located.

Our paper uses a differences-in-differences approach that leverages a wide range of policy shocks to establish a link between lottery sales and SNAP benefits and then estimate the relationship between lottery sales and SNAP benefits. Our main analysis utilizes the recent 2018–2019 US government shutdown, the longest on record, to establish the link between SNAP and lottery sales. The shutdown did not *ex post* change the amount of SNAP benefits that households received, but it did affect how they were disbursed and increased uncertainty around the program. We then use the American Reinvestment and Recovery Act of 2009 (ARRA) to estimate the elasticity of lottery ticket sales with respect to SNAP benefits because this legislation generated changes in the benefits received by households.

We find that during the 2018–2019 government shutdown, sales of scratch-off tickets were 7% lower at SNAP-eligible stores than they would have been if they had not been affected by the shutdown, while sales of draw tickets (e.g., Powerball) were unaffected. This finding demonstrates a link between SNAP and lottery sales. Since SNAP benefits were unchanged *ex post*, we use other policy shocks to estimate the elasticity of lottery sales with respect to SNAP benefits. Using the introduction of ARRA in which SNAP benefits increased by an average of 13.6%, we find that sales of draw lottery tickets declined by 28% relative to their pre-ARRA trend, implying an elasticity of -2.1. Upon ARRA's expiration, the 5% drop in SNAP benefits precipitated a 2% increase in draw lottery sales and a 7% increase in scratch-off sales relative to pre-expiration trends. These estimates imply an elasticity of -1.4 to -0.4. While these estimates range widely, they all suggest that SNAP benefits actually discourage lottery gambling. One possible mechanism for this decrease is that increases in SNAP benefits decrease shopping frequency (Makioka 2018), and therefore reduce the opportunities for lottery gambling.

The rest of the paper is structured as follows. Section 3.2 summarizes important details about the SNAP program. Section 3.3 describes the data. Section 3.4 presents facts about household lottery spending. Section 3.5 presents the model and estimation results and Section 3.6 concludes.

3.2 Background

Supplemental Nutrition Assistance Program (SNAP)

The Supplemental Nutrition Assistance Program (SNAP) is the largest federal nutritional assistance program. It is administered by Food and Nutrition Service (FNS) of the United States Department of Agriculture (USDA) and provides assistance for low-income households to purchase food. It was previously known as the "food stamp" program. SNAP benefits can be used to buy food that will be eaten at home such as breads, cereals, fruits, vegetables, meat, and dairy products. SNAP benefits cannot be used to purchase alcohol, tobacco, health items, personal care items, soap, paper products, household supplies, or hot foods.

Household Eligibility and Benefits

To be eligible for the program, households must have a gross income below 130% of the Federal poverty level.[58] Additionally, households must meet various work requirements and cannot have more than $2,250 in assets, excluding housing and vehicles. Upon being approved, households receive an Electronic Benefit Transfer (EBT) card with their SNAP benefits loaded onto it. The EBT card acts as a debit card for SNAP-eligible food items at stores that accept SNAP benefits. SNAP benefits are disbursed monthly to eligible households.

Benefit amounts decrease as household income increases. As an example, in Pennsylvania in 2019, the maximum benefit a 4-person household could receive is $646. However, the actual benefit received is this maximum amount minus 30% of the household's net monthly income. Therefore, if the household received $1,000 in monthly net income, it would only have a SNAP benefit of $646 - $300 = $346. As of February 2018, 1.8 million people were receiving SNAP benefits in Pennsylvania (*SNAPFacts*).

Store Eligibility

Stores must apply to the USDA to be eligible to accept SNAP benefits. Only stores that sell "staple foods" are eligible to apply. "Staple foods" include fruits and vegetables; meat, poultry, and fish; dairy products; and breads or cereals. Stores where at least 50% of their sales are from "staple foods" will be approved. If stores do not meet the sales requirement, as long as they maintain a minimum level and variety of "staple food" inventory, they will also be approved. According to the USDA, most stores are approved based on meeting the minimum inventory requirements. A summary of additional regulatory detail is in Appendix 3.A.

Policy Changes Affecting SNAP

American Recovery and Reinvestment Act of 2009

On February 17, 2009, the American Recovery and Reinvestment Act (ARRA) was passed. This stimulus bill increased household SNAP benefits by 13.6% (effective April

[58] Residents of Alaska and Hawaii have higher limits compared to the contiguous United States.

1, 2009), which was an average of \$80 per month for a household of 4. Additionally, it expanded eligibility of jobless adults without children. The benefit increase expired on November 1, 2013, which decreased benefits by an average of 5% (Beatty and Tuttle 2014; Rosenbaum and Keith-Jennings 2013).

U.S. Government Shutdown (2018-2019)

The 2018-2019 government shutdown was the longest in US history, lasting 35 days from December 22, 2018, through January 25, 2019. The shutdown was precipitated by disagreements over funding for a wall along the US-Mexico border and affected the SNAP program. To ensure the food security of SNAP recipients, the USDA worked with states to issue February SNAP benefits on or before January 20, 2019 (USDA 2019). In Pennsylvania, the February benefits were issued statewide on January 16, 2019, which was different than the usual disbursement schedule in which SNAP benefits are issued between the first and tenth day of each month (Lubrano 2019; USDA 2020).

Pennsylvania Lottery

Created in 1971, the Pennsylvania Lottery is the state-run lottery of Pennsylvania and is the only state lottery to use its proceeds to fund programs for older residents such as centers, meals, prescription assistance, and transportation. It offers many different games, which can be grouped into two main categories:

- **Draw Games:** Players choose a set of two to six numbers (depending on the specific game). The Lottery randomly draws a set of two to six numbers as well and if players match a subset, they win a fixed prize. If they match all numbers drawn, they win the jackpot, which continues to accumulate value until someone wins it. Common examples of draw games are Powerball and Mega Millions. Ticket prices range from \$0.50 to \$3.
- **Instant Games:** These games are also known as "scratchers" or "scratch-offs." Players purchase a ticket with spaces that can be scratched off to reveal symbols. If those symbols match various winning combinations, the player wins a prize based on which combination is matched. Prizes and odds are pre-specified and

do not change, in contrast to draw games where the jackpot can vary over time. Prizes range from $5 to $3 million and ticket prices range from $1 to $30. More expensive tickets offer better odds and/or larger prizes.[59]

Any Pennsylvania business can become a licensed lottery retailer as long as it passes a credit check, background investigation, tax clearance check, and is compliant with the Americans with Disabilities Act.[60] Lottery retailers earn commissions on the lottery tickets that they sell as well as other bonuses and incentives.

3.3 Data

We obtain data from two administrative sources: (1) the Pennsylvania Department of Revenue provides store-level lottery ticket sales by game and month and (2) the USDA's Food and Nutrition Service (which administers SNAP) provides store-level eligibility spells to accept SNAP benefits. This section briefly describes the data and how they were combined.

SNAP Store Eligibility Data

We obtain store eligibility data from the USDA's Food and Nutrition Service. This data contains the name and address of each SNAP-eligible store, the store type, and their dates of eligibility (i.e., an observation is a store-eligibility period). The USDA specifies 17 types of stores, which we aggregate into three categories: convenience stores (including small, urban grocery stores), supermarkets, and restaurants.[61] We geocode each store's location using Texas A&M's Geocoding Services.[62] We restrict our analysis to addresses that were able to be matched exactly based on the address or parcel or interpolated based on the address range ("high-quality" matches). For

[59]The authors stress that the expected value of any lottery ticket remains negative, it just might be less negative for more expensive tickets. Buying a lottery ticket remains a poor investment.

[60]Retailers that are not compliant can get assistance from the Pennsylvania Lottery to offset costs incurred to become compliant through the American with Disabilities Act PA Lottery ADA Program.

[61]Restaurants are generally not eligible to accept SNAP benefits, but some neighborhood delis also offer a selection of food items like those seen in convenience stores, which could meet the USDA eligibility requirements.

[62]Available at https://geoservices.tamu.edu/

example, if the geocoding database does not contain 131 Main Street, but has geocodes for 123 Main Street and 141 Main Street, it will interpolate the geocoded location between those addresses.

After geocoding, we have 15,377 unique locations and 95% of these locations are "high-quality" matches. Locations may have more than one eligibility spell, but we find that most do not. 65% of locations only have one eligibility spell and 19% have two spells. Finally, locations tend to be eligible for long periods of time with the median eligibility spell being 5 years and the interquartile range is from 2 to 13 years. After organizing, this data consists of SNAP-eligibility spells by geocoded location.

Pennsylvania Lottery Data

Monthly store-level lottery ticket sales from October 2002 through March 2019 were obtained through an open-records request to the Pennsylvania Department of Revenue.[63] The data consists of names and addresses of each lottery terminal as well as the monthly net sales of lottery tickets by individual ticket type. Since retailers can have multiple terminals, we aggregate all sales to the geocode level. For each location, we aggregate lottery ticket sales into two categories: instant sales (e.g., "scratchers") and draw sales (e.g., Powerball).[64]

Over the June 2018 to March 2019 time period, we observe 14,255 unique lottery machines across 9,335 unique locations.[65] The vast majority of locations have only one or two lottery machines. Each month, about 13,000 machines record positive sales. The distribution of monthly ticket sales is reported in Table 3.1. The table shows that scratch-off tickets account for about two-thirds of total lottery sales.

In order to match this data with the SNAP data, we geocode each store's location using the same approach as outlined for the SNAP data. 86–88% of lottery ticket sales can be assigned to high-quality geocodes and 87% of lottery machines with positive sales are captured.

[63] Currently we are awaiting updated data from October 2002–May 2018 because of coverage issues identified by authors. For comparison, the earlier data only captured 20% of publicly reported sales while the June 2018 data was within 2% of publicly reported sales. Many high-grossing stores are not present in the earlier data sample.

[64] We also obtained jackpot data for Powerball and MegaMillions from BeatLottery.net.

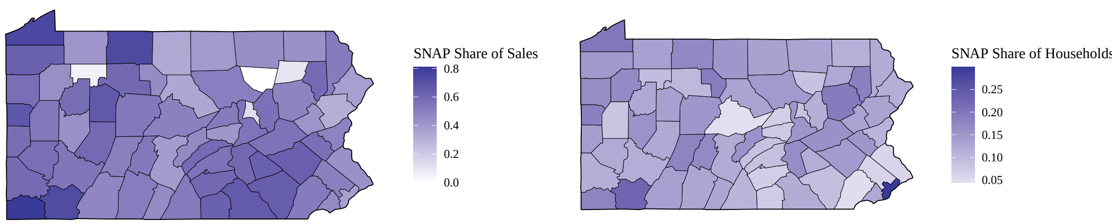
[65] The number of geographic locations is likely underestimated because in unmatched cases, the geocoder assigns the location to the centroid of the ZIP code.

Table 3.1: Distribution of Monthly Lottery Ticket Sales

Variable	Mean	SD	25th Pctile	Median	75th Pctile
Total	29,679	34,652	5,347	16,526	43,205
Scratch-Off	20,367	25,287	2,100	10,500	30,466
Draw	9,312	13,382	987	4,551	12,138

Note: Using lottery sales data from June 2018 through March 2019, this table reports the distribution of monthly lottery sales.

Figure 3.1: Lottery Sales and SNAP Participation by County



(a) SNAP Share of Lottery Sales (b) County SNAP Participation

Note: Using PA Lottery sales data and SNAP store-eligibility, Figure (a) plots the share of lottery ticket sales at SNAP-eligible stores by county and Figure (b) plots the share of SNAP households by county.

Merging Datasets

After geocoding each location, we match each entry based on their geocodes (i.e., latitude and longitude). If a lottery retailer and a SNAP-eligible location have the same geocode, we treat them as the same location.[66] Since lottery retailers that are not in the SNAP data do not have store types, we manually classify many of these stores. The summary of the store types is provided in Table 3.2. While SNAP-eligible

[66]We make some manual corrections because stores in shopping plazas often share the same geocode. In these cases, there may be different store types (e.g., a grocery store and a liquor store) at the same geocode. This matching strategy is better than a fuzzy name and address match because lottery store names are recorded by location, but companies with multiple stores may only file a single SNAP application under the holding company name. As a result, if ABC Holdings, LLC owns Shop XYZ, "Shop XYZ" will be recorded in the lottery data while "ABC Holdings, LLC" will be recorded in the SNAP data.

stores account for 41% of lottery retailers, they make up 52% of lottery sales, a disproportionately higher share. Furthermore, the types of stores that accept SNAP benefits are primarily grocery and convenience stores. Most large grocery stores are eligible to accept SNAP benefits while convenience stores are more evenly split between those that accept SNAP benefits and those that do not. The remainder of non-SNAP lottery retailers are primarily bars, restaurants, check-cashing establishments, and tobacco shops.[67]

Table 3.2: Lottery Sale Shares by Store Type

	Stores		Sales	
Store Type	SNAP	Non-SNAP	SNAP	Non-SNAP
Conv. Store	0.32	0.28	0.39	0.30
Grocery	0.08	0.01	0.14	0.01
Food / Alcohol	0.005	0.14	0.002	0.06
Misc	0.001	0.16	0.0005	0.11
Total	0.41	0.59	0.52	0.48

Note: Using lottery sales data from June 2018 through March 2019, this table reports the share of lottery sales at SNAP and non-SNAP-eligible retailers. “Conv. Stores” denotes gas stations, convenience stores, corner shops, and pharmacies. “Grocery" denotes grocery stores that offer a large selection of fresh produce and other food items. “Food / Alcohol” denotes restaurants, bars, and beer/wine/spirits stores. “Misc” denotes check-cashing stores, social clubs, and remaining unclassified lottery retailers.

Consumer Expenditure Survey

We also use the Consumer Expenditure Survey (CEX) from the Bureau of Labor Statistics to obtain information on household-level gambling. The CEX samples about 7,000 households and interviews them for four consecutive quarters. Detailed household demographics and income information are collected at the start and end of a household’s tenure in the sample.

[67] The “Misc” category includes a long tail of establishments that have not been classified and may include stores that fit into other categories. Based on inspection, many of these appear to be convenience stores that were not classified based on our existing classification algorithm.

3.4 Lottery Gambling Patterns

This section examines aggregate patterns in lottery gambling. First, we present aggregate county-level patterns and then we examine household-level patterns in lottery spending.

To examine county-level patterns in lottery gambling, we regress per-capita lottery sales on county-level characteristics, including income, unemployment, and a range of government transfer payments, including SNAP payments.

$$\log(Sales_{ct}) = \alpha + \beta X_{ct} + \lambda_c + \lambda_t + \epsilon_{ct}, \tag{3.1}$$

where $Sales_{ct}$ are total lottery ticket sales in county c in year t. X includes log per-capita measures of income, unemployment insurance, Social Security payments, SNAP payments, and the Earned Income Tax Credit payments, along with the unemployment rate. λ captures county and year fixed effects. We estimate this regression using lottery sales from 2003–2017. Table 3.3 reports the results for draw and scratch-off sales separately.

Across both draw and scratch-off sales, increases in county-level SNAP payments are associated with increases in lottery sales, while increases in the unemployment rate are associated with decreases in lottery sales. Counties that experience a 10% increase in per-capita SNAP payments are expected to increase lottery sales by 1.4–1.9%. While this reveals aggregate patterns in lottery ticket sales, it is unclear what is driving this pattern, assuming SNAP transfers directly affect lottery sales.

Increases in SNAP transfers could be generated by more households obtaining SNAP (extensive margin) or by existing SNAP households receiving more generous benefits (intensive margin). If increases in SNAP transfers are primarily generated by the extensive margin, then lottery sales could increase if SNAP households are more likely to play the lottery. On the other hand, if increases in SNAP transfers are primarily generated by increased benefits to existing households, then lottery sales could increase if these benefits are "passed through" to purchase lottery tickets. In order to determine whether there are significant differences in gambling between SNAP and non-SNAP households, we turn to the Consumer Expenditure Survey (CEX).

The CEX is advantageous for this purpose because since 2004, households have been asked about both their lottery consumption and their SNAP benefits. In particular,

Table 3.3: County Lottery Sales and Government Transfers

	Log Sales	
	Draw	Scratch
	(1)	(2)
Log(Income)	0.085	0.715***
	(0.145)	(0.157)
Log(UI)	0.092	0.035
	(0.061)	(0.066)
Log(Soc. Sec.)	0.671***	0.411
	(0.258)	(0.279)
Log(SNAP)	0.141**	0.194***
	(0.064)	(0.070)
Log(EITC)	0.053	0.003
	(0.137)	(0.148)
UR	−0.025*	−0.033**
	(0.015)	(0.016)
Avg Sales	25.39	54.62
Year FE's	Y	Y
County FE's	Y	Y
Observations	990	990
Adjusted R^2	0.893	0.861
Note:	*p<0.1; **p<0.05; ***p<0.01	

Note: Using lottery sales data from 2003–2017 combined with BEA's Regional Economic Accounts data, this table reports the results of regressing log annual lottery ticket sales, by county, on county-level transfer payments. All variables are log per-capita measures except for the unemployment rate, which is included as is.

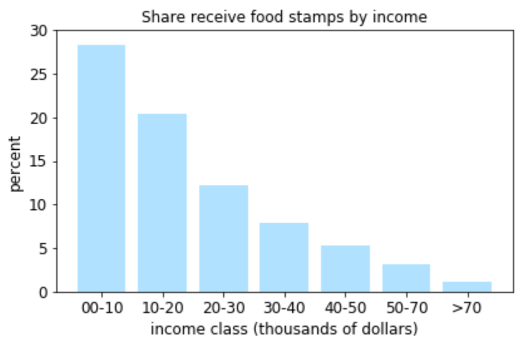
households report the dollar amount consumed each quarter on "lotteries and games of chance" as well as the dollar value of the last food stamps or EBT received each quarter. The former offers intensive and extensive margins for analyzing lottery purchases between SNAP and non-SNAP households, a distinction which the latter provides. Although the CEX is a (representative) survey of household expenditures, it is better used as a cross-section of households instead of a panel to track changes within households, meaning this dataset can at most afford us a correlation between SNAP benefits and lottery expenditures.

Panels (a) and (b) of Figure 3.2 summarize SNAP benefits across income groups in the CEX dataset. Reassuringly, SNAP recipients in the CEX are predominantly concentrated among the lowest income brackets. Across SNAP-receiving households, we see less variation in the dollar amount that SNAP households on average receive. Households that earn less than \$10,000 per year in total income earn about \$800 per quarter on average from SNAP benefits, where as SNAP-eligible households above \$10,000 accumulate roughly \$1,200 per quarter in benefits. How does this translate to expenditures on lotteries? Panel (c) reveals that SNAP-eligible households across the income spectrum purchase lottery tickets at similar rates to non-SNAP households, and perhaps at even higher rates within higher-income brackets.[68] This confirms that SNAP-eligible households spend money on lotteries. Panel (d) reveals that SNAP-households spend a similar, perhaps slightly smaller, amount of their quarterly income on lotteries compared with non-SNAP households. Panels (e) and (f) investigate if there is a relation between the magnitude of SNAP benefits and lottery purchases: There does appear to be a slight uptick in lottery incidence for households receiving more SNAP benefits, but no clear trend is apparent in lottery expenditures.

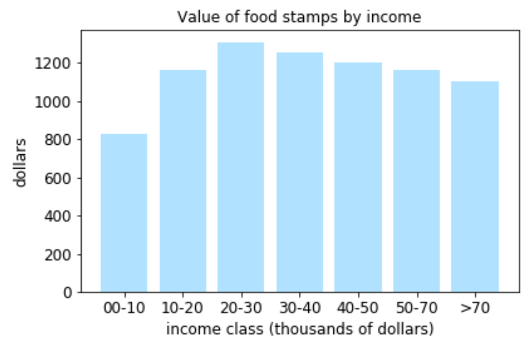
To investigate the relation between lotteries and SNAP benefits further, we turn to a regression analysis. The first four columns of Table 3.4 explore the the extensive margin of lottery purchases by implementing a linear probability model. The last four columns explore the intensive margin of lottery purchases, conditional on households purchasing lotteries that year-quarter. SNAP households in the CEX data are 1.7

[68]Although are incidence rates for lottery expenditures is relatively low compared with other studies—which may be due to the miscellaneous nature of reporting lottery expenditures in the CEX survey—reassuringly we find a similar pattern of increasing incidence with income as in Kearney (2005).

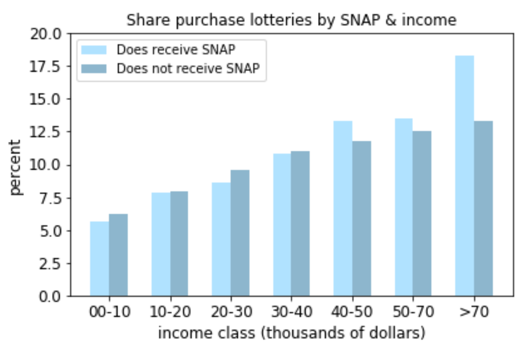
Figure 3.2: SNAP and Lotteries in CEX



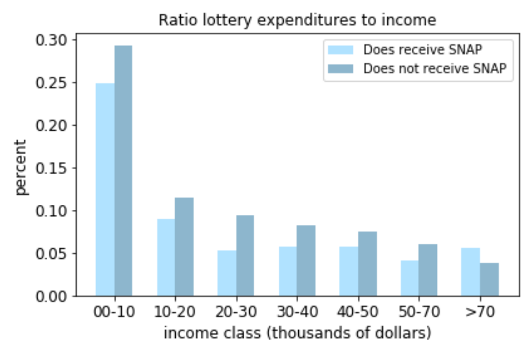
(a) Snap incidence



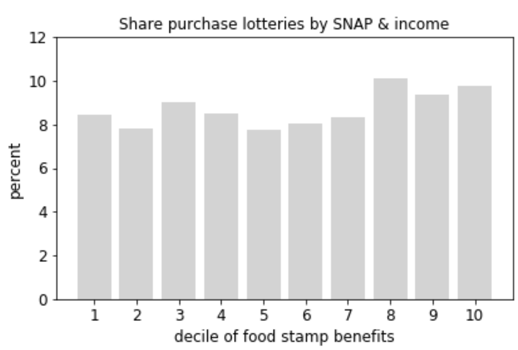
(b) Snap value



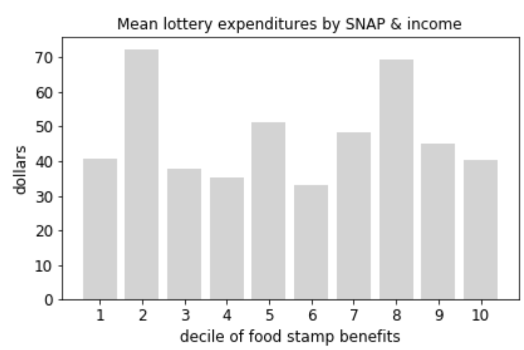
(c) Lottery incidence



(d) Lottery expenditure to income ratio



(e) Lottery incidence for recipients



(f) Lottery expenditures for recipients

Note: CEX Data

Table 3.4: CEX - Purchases Lotteries

	Has lottery purchase				Log lottery purchase			
Receives food stamps	-0.017***	-0.002			-0.307***	-0.169***		
	(0.004)	(0.003)			(0.045)	(0.045)		
Log value of food stamps			0.006**	0.002			0.021	-0.003
			(0.003)	(0.003)			(0.022)	(0.020)
Log family income		0.017***		0.025***		0.164***		0.161***
		(0.001)		(0.002)		(0.010)		(0.035)
Education and race FE, age cubic	✓	✓	✓	✓	✓	✓	✓	✓
State and year-quarter FE	✓	✓	✓	✓	✓	✓	✓	✓
N	321269	321269	26190	26190	37029	37029	2362	2362
Adjusted R^2	0.03	0.03	0.02	0.03	0.10	0.11	0.10	0.11

Notes: CEX data.

percentage points less likely to purchase lotteries compared with non-SNAP households (Column 1), however this gap is driven by differences in household income rather than preferences (Column 2). Among SNAP-eligible households, while there appears to be a positive relation between SNAP benefit amount and lottery expenditures (Column 3), this slope appears to be driven by differences in total household income rather than SNAP income.

While SNAP households purchase lotteries at similar rates conditional on their total income, they do appear to spend less on lotteries. On average, SNAP households spend about 26% less (31 log-points) on lotteries than non-SNAP households (Column 5). Controlling for household income (including SNAP benefits) reduces this gap to 16% between SNAP and non-SNAP households. And conditional on households receiving SNAP benefits, there does not appear to be any relation between the amount of benefits a household earns and its lottery expenditures (Columns 7 and 8). Importantly though, this only reveals that higher-SNAP-benefit households do not purchase more lottery tickets than lower-SNAP-benefit households, not necessarily that a single SNAP-eligible household that experienced a windfall (haircut) in SNAP benefits would buy more (less) lottery tickets.

3.5 Empirical Model

The previous section found no evidence that SNAP and non-SNAP households play the lottery at different rates. If anything, SNAP households spend less on lotteries than non-SNAP households. This section examines whether changes in SNAP benefits (intensive margin) affect lottery ticket sales. To establish a link between SNAP benefits and lottery sales, we use a differences-in-differences approach that leverages shocks to the SNAP program, but that did not affect the Pennsylvania Lottery.

Government Shutdown

The first shock we leverage is the 2018–2019 government shutdown. During the shutdown, there was heightened uncertainty about whether SNAP benefits would be affected. Ultimately, the timing of benefit payouts was modified, but the total amount of benefits was unchanged. If a relationship exists, then the early SNAP disbursement will only affect SNAP-eligible lottery retailers and SNAP-ineligible retailers will serve as the control group.

A priori household behavior in response to these "early, not extra" benefits is unclear. On one hand, the USDA and state agencies had limited time and resources to communicate this change to recipients and households may have treated these benefits as extra and therefore increased their spending. On the other hand, households may be aware that these are "early, not extra" and appropriately budgeted, therefore spending like they normally would. On the third hand, the shutdown increased uncertainty because households were suspicious that the government would penalize them for using more benefits than they were entitled to (Lubrano 2019). Additionally, the shutdown duration was indeterminate and there was a risk that March benefits would be reduced and April benefits may be eliminated if the shutdown extended into the spring (Lubrano 2019). Because of this uncertainty and the fact that SNAP benefits roll over each month, households may have reduced spending to generate precautionary savings against an extended shutdown. Regardless of how households treated these benefits, the important factor is that this federal shock is not expected to affect the state-run lottery.

The differences-in-differences specification is as follows:

$$\log(Sales_{ict}) = \alpha + \sum_{\tau \neq \text{Dec 2018}} \beta_\tau SNAP_i * \mathbb{1}\{t = \tau\} + \lambda_{ct} + \lambda_i + \epsilon_{ict}, \tag{3.2}$$

where $Sales_{ict}$ denotes the lottery ticket sales of store i in county c month t. $SNAP_i$ is an indicator for the SNAP-eligibility of store i. λ_i and λ_{ct} are store and county-month-year fixed effects, respectively.

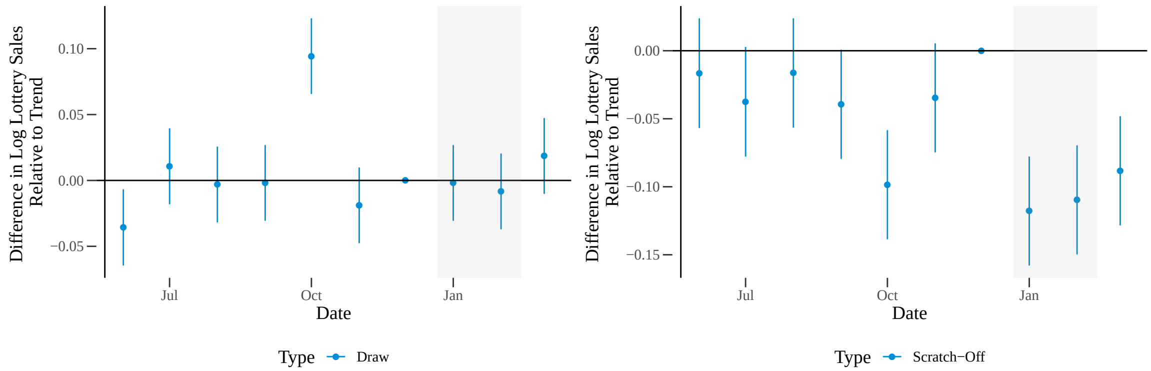
This specification models how lottery sales evolve over time and β_t captures whether lottery sales at SNAP-eligible stores behave differently than at non-SNAP-eligible stores in month t (December 2018 is the reference month). We remove any stores that change their SNAP eligibility status over this period to ensure that β_t is identified off of changes in lottery sales over time between the two groups and not changes in the SNAP eligibility of a store. Even though lottery sales are highly seasonal with promotions around holidays, the non-SNAP month indicators capture this seasonality, so we can identify the effect of the government shutdown separate from seasonal fluctuations in sales from β_τ.

As a further check, if lottery sales at SNAP-eligible stores behave differently than sales at non-SNAP stores, then the effect would also be tied to the density of households receiving SNAP benefits in the surrounding county. SNAP-eligible stores in counties with low densities of households that participate in SNAP may experience less of a shock than SNAP-eligible stores in counties with high densities of SNAP-participating households. To test this, we replace the SNAP dummy variable with an interaction term of the SNAP dummy and county-level SNAP participation rates. Any county-specific trends will be captured by the county-year-month fixed effect.

Figure 3.3 plots the values of β_t and illustrates that there are no significant differences in log lottery sales trends prior to January 2019, with the exception of a spike in October 2018. This spike is likely attributable to the fact that both the Powerball and Mega Millions jackpots were at record or near-record highs of \$602.5 million and \$1.6 billion, respectively. Figure 3.4 illustrates that draw sales track jackpots while instant sales are less responsive.

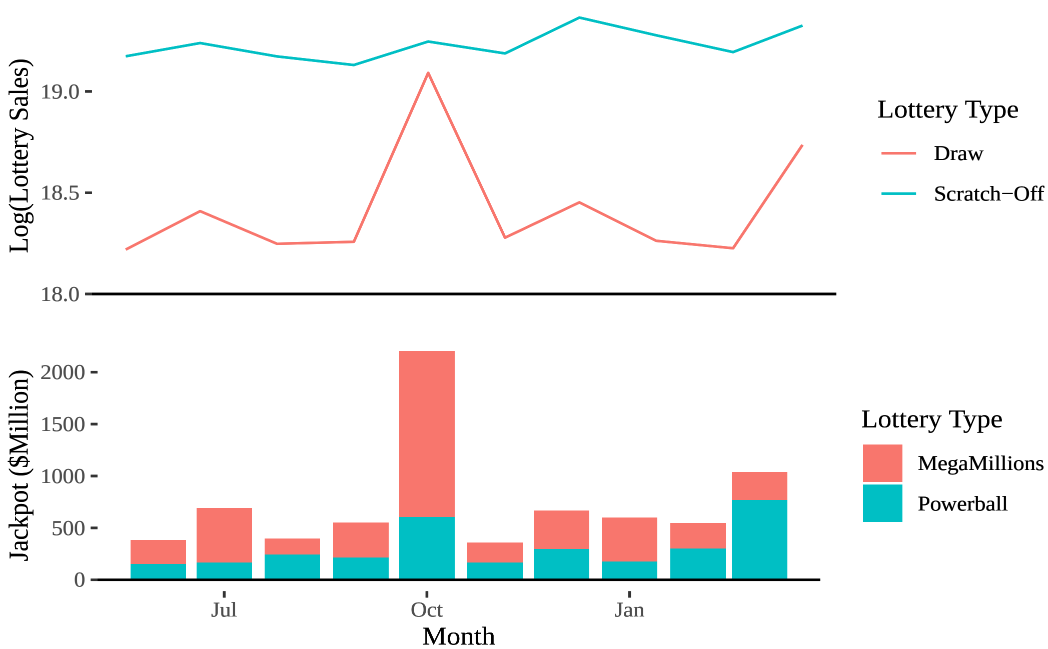
During the shutdown, scratch-off sales at SNAP stores decreased by a dramatic 12% and then started recovering through March, but were still depressed by 9% relative

Figure 3.3: SNAP-Eligible Store Lottery Sale Changes



Note: Using PA Lottery sales data and SNAP store-eligibility, this figure plots differences in lottery ticket sales at SNAP-eligible relative to expected sales had they evolved similarly to sales at non-SNAP-eligible stores by month. Shading denotes period where SNAP benefits were disbursed early (i.e., January disbursements were doubled and February disbursements were zero).

Figure 3.4: Relationship of Lottery Sales and Jackpot Sizes



Note: Using PA Lottery sales data and jackpot data, this figure plots monthly lottery ticket sales by game type and the monthly maximum jackpot amount for the Powerball and Mega Millions multi-state lotteries.

to trends at non-SNAP stores. On the other hand, draw sales were not significantly different between SNAP and non-SNAP eligible stores.

We should note that these estimates may be understated. The main reason for this is that households could "pass-through" their SNAP benefits to lottery tickets at non-SNAP-eligible stores. While we do not assume that all SNAP pass-through is spent on lottery tickets at SNAP-eligible stores, we believe that it is more likely to occur at a SNAP-eligible store due to convenience. Because this bias works against us, our estimates provide a lower bound on the magnitude of the shutdown's effect on lottery sales.

Finally, to concisely report the estimation results for the period of the shutdown, we estimate a modified version of Equation 3.2 that better focuses on the shutdown period:

$$\log(Sales_{ict}) = \alpha + \sum_t \beta_t SNAP_i * \mathbb{1}\{t \in Shutdown\} + \lambda_{ct} + \lambda_i + \epsilon_{ict}, \tag{3.3}$$

where $Sales_{ict}$ denotes the lottery ticket sales of store i in county c in month t. $SNAP_i$ is an indicator for the SNAP-eligibility of store i. $\mathbb{1}\{t \in Shutdown\}$ indicates if month t is during the impacted period of January or February 2019. λ_i and λ_{ct} are store and county-month-year fixed effects, respectively.

Table 3.5 reports the results. Columns (1) and (2) show that draw lottery sales were not significantly affected during the shutdown while scratch-off sales decreased by 7–8% during the months affected by the shutdown. Columns (3) and (4) show that this effect is stronger at SNAP-eligible stores in counties with higher levels of SNAP participation. A store in a county with a ten percentage point higher SNAP participation rate would be expected to experience 1.4 percentage points lower draw sales during January and 3.7 percentage points lower scratch-off sales.

American Recovery and Reinvestment Act of 2009

In order to quantify the relationship between SNAP benefits and lottery sales, we use the passage of the American Recovery and Reinvestment Act (ARRA). ARRA increased the SNAP benefits for all SNAP recipients by an average of 13.6%, effective April 1, 2009. We use this variation to estimate another differences-in-differences

Table 3.5: SNAP Lottery Sales Estimation Results (Gov't Shutdown)

	Log Sales			
	Draw	Scratch	Draw	Scratch
	(1)	(2)	(3)	(4)
Jan-19	-0.009	-0.075^{***}		
	(0.012)	(0.016)		
Feb-19	-0.017	-0.067^{***}		
	(0.012)	(0.016)		
Jan-19 : SNAP Density			-0.144^{*}	-0.367^{***}
			(0.074)	(0.100)
Feb-19 : SNAP Density			-0.042	-0.360^{***}
			(0.074)	(0.100)
Avg Sales	12512	27091	12512	27091
Store FE's	Y	Y	Y	Y
County-Month FE's	Y	Y	Y	Y
Observations	80,990	80,990	80,990	80,990
Adjusted R^2	0.887	0.809	0.887	0.809
Note:	*p<0.1; **p<0.05; ***p<0.01			

Note: Using 2018–2019 PA lottery sales data, this table reports the estimation results of Equation 3.3. "Jan-19" and "Feb-19" indicate the months of SNAP benefits affected by the shutdown. "SNAP Density" denotes the share of households within a county that participate in SNAP, as of July 2018.

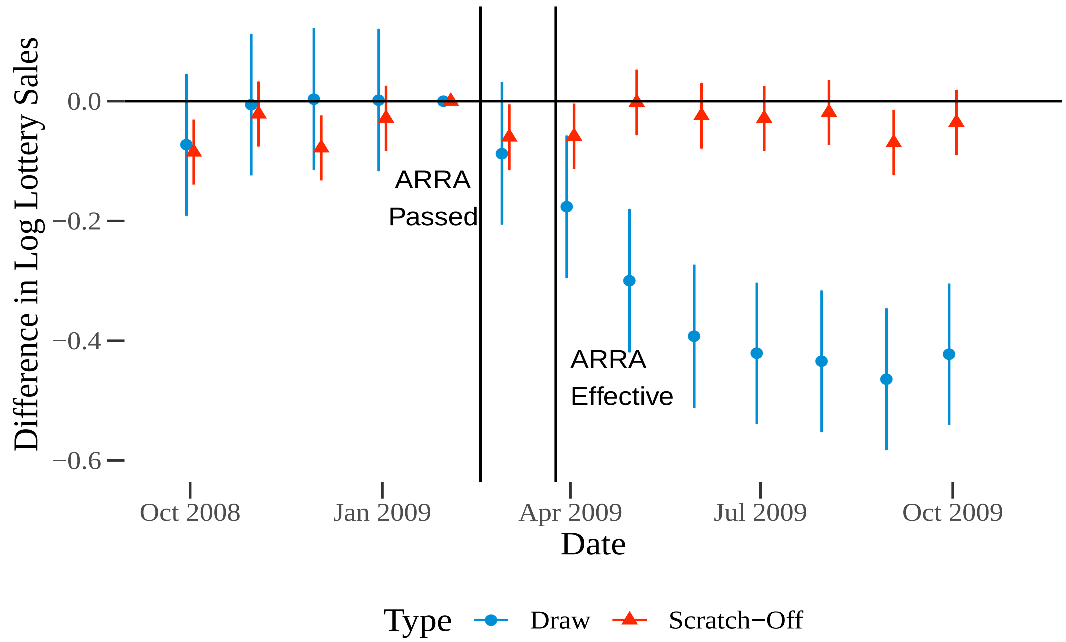
model on lottery sales data from October 2008 through October 2009:[69]

$$\log(Sales_{ict}) = \alpha + \sum_{\tau \neq \text{Feb 2009}} \beta_\tau SNAP_{ict} * \mathbb{1}\{t = \tau\} + \lambda_{ct} + \lambda_i + \epsilon_{ict}, \tag{3.4}$$

where $Sales_{ict}$ denotes the lottery ticket sales of store i in county c in month t. $SNAP_{ict}$ is an indicator for the SNAP-eligibility of store i. λ_i and λ_{ct} are store and county-month-year fixed effects, respectively. We restrict our sample to a balanced panel of stores that maintain the same SNAP status throughout the period and are present for all 13 months time period studied. This restriction helps ensure that β is identified off of changes in lottery sales over time between the two groups and not

[69] This estimation uses the non-universe sample of stores that only captures 20% of all lottery sales. Results may change when the updated data is received.

Figure 3.5: SNAP-Eligible Store Lottery Sale Changes (ARRA Passage)



Note: Using PA Lottery sales data and SNAP store-eligibility, this figure plots difference in lottery ticket sales at SNAP-eligible relative to expected sales had they evolved similarly to sales at non-SNAP-eligible stores by month.

from changes in a store's SNAP eligibility or from possible attrition given that the Great Recession is occurring at this time.

Figure 3.5 plots the β coefficients over time. There is a striking drop in draw lottery ticket sales at SNAP-eligible stores relative to what would have been expected had sales evolved like those at non-SNAP-eligible stores. This drop begins soon after ARRA is passed and then exhibits a significant and sustained decline after ARRA is effective.

In order to obtain the average treatment effect, we estimate Equation 3.5 that compares lottery sales in the pre-ARRA period with lottery sales in the post-ARRA period (post-ARRA refers to its effective date):

$$\log(Sales_{ict}) = \alpha + \beta SNAP_{ict} * PostARRA + \lambda_{ct} + \lambda_i + \epsilon_{ict}, \tag{3.5}$$

Table 3.6 reports the results. Column (1) shows that after ARRA became effective, draw lottery sales at SNAP-eligible stores experienced a 28% decline relative to their expected evolution if they followed the sales pattern at non-SNAP eligible stores. However, scratch-off sales show no significant change after ARRA became effective. Columns (3) and (4) allow for the effect of ARRA to vary based on the share of households that receive SNAP benefits in each county. Counties with larger shares of SNAP households experienced larger declines in draw lottery sales.

Table 3.6: Differences-in-Differences Estimation Results (ARRA Passage)

	Log Sales			
	Draw	Scratch	Draw	Scratch
	(1)	(2)	(3)	(4)
Post-ARRA	-0.331^{***}	0.013		
	(0.023)	(0.010)		
Post-ARRA : SNAP Density			-2.633^{***}	0.052
			(0.163)	(0.069)
Avg Sales	17304	21686	17304	21686
Store FE's	Y	Y	Y	Y
County-Month FE's	Y	Y	Y	Y
Observations	22,412	22,412	22,412	22,412
Adjusted R^2	0.830	0.900	0.831	0.900
Note:	*p<0.1; **p<0.05; ***p<0.01			

Note: Using lottery sales data from October 2008–October 2009, this table reports the differences-in-differences coefficients from estimating Equation 3.5 which regresses log lottery sales on an indicator for post-ARRA after controlling for store and county-year-month fixed effects. Columns (3) and (4) allow for the effect to vary based on county level SNAP participation rates.

By combining the change in SNAP benefits that resulted from ARRA with the change in lottery ticket sales at SNAP-eligible stores, we can quantify how much lottery sales are expected to change as SNAP benefits change. Since the average SNAP benefit increase was 13.6% and the estimated draw lottery sale decrease is 28% (or 33 log-points), this implies that a 1% increase in SNAP benefits will correspond with a 2.1% decrease in draw lottery sales.

This 28% decline is dramatic and should be interpreted with caution. Compared to the universe of lottery sales used in our analysis of the government shutdown, our analysis of ARRA only uses a sample that captures about 20% of lottery sales.

Relative to the universe of stores, the composition of the sample is similar, based on breakdown by store type and SNAP status. However, since we only have a sample of stores, there is a risk this pattern could be generated by households shifting their gambling from stores in our sample to stores not in our sample and that this is more common for lottery purchases at SNAP stores. We will be able to address this concern when we receive the universe of lottery sales because these shifts will show up as increased sales at other stores.

Expiration of ARRA Benefits

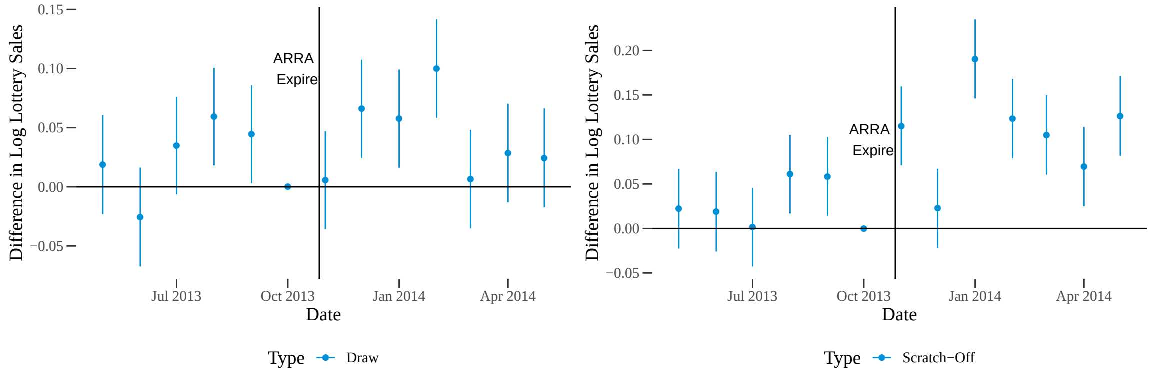
Not only did ARRA increase benefits when it was adopted, but it also reduced benefits when it expired in 2013. Upon expiration, SNAP benefits decreased by 5%, effective November 1, 2013. This expiration provides another opportunity to examine how lottery sales respond to changes in the SNAP program. We use this variation to estimate another differences-in-differences model using data from May 2013 to May 2014:

$$\log(Sales_{ict}) = \alpha + \sum_{\tau \neq \text{Oct 2013}} \beta_\tau SNAP_{ict} * \mathbb{1}\{t = \tau\} + \lambda_t + \lambda_{ic} + \epsilon_{ict}, \tag{3.6}$$

where $Sales_{ict}$ denotes the lottery ticket sales of store i in county c in month t. $SNAP_{ict}$ is an indicator for the SNAP-eligibility of store i. λ_i and λ_{ct} are store and county-month-year fixed effects, respectively. We restrict our sample to a balanced panel of stores that maintain the same SNAP status throughout the period and are present for all 13 months time period studied. This restriction helps ensure that β is identified off of changes in lottery sales over time between the two groups and not from changes in a store's SNAP eligibility or from possible attrition.

Figure 3.6 plots the β coefficients over time. There appears to be a temporary increase in draw lottery sales for about three months after ARRA expires and then draw lottery sales revert back to the same trend as at non-SNAP stores (relative to October 2013) by March 2014. On the other hand, scratch-off sales are elevated after ARRA expires relative to how sales evolved at non-SNAP eligible stores starting in October 2013. The fact that scratch-off tickets exhibit a strong response after ARRA expiration could be due to its increasing popularity since 2010. Between 2010 and 2014, draw sales declined from 43% to 36% of total lottery sales. This is primarily

Figure 3.6: SNAP-Eligible Store Lottery Sale Changes (ARRA Expiration)



Note: Using PA Lottery sales data and SNAP store-eligibility, this figure plots differences in lottery ticket sales at SNAP-eligible relative to expected sales had they evolved similarly to sales at non-SNAP-eligible stores by month.

because levels of draw ticket sales have been relatively flat while scratch-off sales have been the source of almost all growth in lottery revenues. In nominal terms, scratch off sales grew from \$1.7 billion in 2010 to \$2.4 billion in 2014 while draw tickets have inched up from \$1.3 billion in 2010 to \$1.4 billion in 2014.

In order to obtain the average treatment effect, to quantify the relationship between SNAP benefits and lottery sales, we estimate Equation 3.7 that compares lottery sales in the six months prior to ARRA's expiration with lottery sales in the six month period after ARRA's expiration:

$$\log(Sales_{ict}) = \alpha + \beta SNAP_{ict} * ARRAExpire + \lambda_{ct} + \lambda_i + \epsilon_{ict}, \tag{3.7}$$

Table 3.7 reports the results. Column (1) shows that draw lottery sales increased by about two percent after ARRA expired relative to how they evolved at non-SNAP stores. However, Column (2) shows that scratch-off tickets exhibited a larger increase of seven percent after ARRA expired. As before, Columns (3) and (4) show that these effects are increasing with the share of households within a county that receive SNAP benefits. Combining these changes with the average 5% decline in SNAP benefits experienced after ARRA's expiration, we obtain that the elasticity of lottery sales with respect to SNAP benefits is between -0.4 and -1.4. This is a smaller magnitude than was estimated upon ARRA's adoption, but the change in SNAP benefits was smaller and the change was publicized about three months before it would take effect, so households may have been more able to anticipate the change (Rosenbaum and Keith-Jennings 2013).[70]

Table 3.7: Differences-in-Differences Estimation Results (ARRA Expiration)

	Log Sales			
	Draw	Scratch	Draw	Scratch
	(1)	(2)	(3)	(4)
Post-ARRA	0.019***	0.070***		
	(0.006)	(0.009)		
Post-ARRA : SNAP Density			0.114***	0.450***
			(0.040)	(0.058)
Avg Sales	16516	31218	16516	31218
Store FE's	Y	Y	Y	Y
County-Month FE's	Y	Y	Y	Y
Observations	21,294	21,294	21,294	21,294
Adjusted R^2	0.977	0.920	0.977	0.920
Note:	*p<0.1; **p<0.05; ***p<0.01			

3.6 Conclusion and Future Research

This paper presents a new finding that SNAP benefits decrease lottery gambling. Using new data on store-level lottery sales in Pennsylvania along with each store's SNAP eligibility status, we have a detailed view of lottery sales at a vast array of retailers. By combining our detailed store panel with policy shocks to the SNAP program, including the 2018–2019 government shutdown and the American Recovery and Reinvestment Act of 2009, we estimate the elasticity of lottery sales with respect to SNAP benefits. While our estimates range widely from -2.1 to -0.4, we find that lottery gambling decreases when SNAP benefits are increased.

These findings prompt a variety of questions related to public finance. Are these findings unique to the SNAP program or do they apply to other transfer programs like

[70] These estimates are subject to the same caveat of sample selection as in the ARRA passage section. However, given that there is not a dramatic economic upheaval during this period, dramatic shifts in consumer behavior are less likely. As before, this possibility will be addressed when we receive the updated universe of data.

unemployment insurance or Temporary Assistance for Needy Families. If the primary mechanism is through shopping frequency, then we are unlikely to see similar effects from non-nutritional assistance programs. More research will be needed to determine what households spend this extra money on. Given that increases in gambling are funded by decreases in food and housing spending, it might be true that savings from *not* gambling (or not having gambling opportunities) are used to increase food and housing expenditures.

References

Beatty, Timothy KM and Charlotte J Tuttle (2014). "Expenditure response to increases in in-kind transfers: Evidence from the Supplemental Nutrition Assistance Program". In: *American Journal of Agricultural Economics* 97.2, pp. 390–404.

Blalock, Garrick, David R. Just, and Daniel H. Simon (2007). "Hitting the Jackpot or Hitting the Skids: Entertainment, Poverty, and the Demand for State Lotteries". In: *American Journal of Economics and Sociology* 66.3, pp. 545–570.

Combs, Kathryn L., Jaebeom Kim, and John A. Spry (2008). "The relative regressivity of seven lottery games". In: *Applied Economics* 40.1, pp. 35–39.

Fu, Wei and Feng Liu (2019). "Unemployment insurance and cigarette smoking". In: *Journal of health economics* 63, pp. 34–51.

Haisley, Emily, Romel Mostafa, and George Loewenstein (2008). "Subjective relative income and lottery ticket purchases". In: *Journal of Behavioral Decision Making* 21.3, pp. 283–295.

Hastings, Justine and Jesse M Shapiro (2018). "How are SNAP benefits spent? Evidence from a retail panel". In: *American Economic Review* 108.12, pp. 3493–3540.

Kearney, Melissa Schettini (2005). "State lotteries and consumer behavior". In: *Journal of Public Economics* 89.11-12, pp. 2269–2299.

Lantis, Robert and Brittany Teahan (2018). "The effect of unemployment insurance on alcohol use and abuse following job loss". In: *Economics & Human Biology* 30, pp. 92–103.

Lubrano, Alfred (Jan. 16, 2019). "Because of shutdown, February food stamps disbursed on Wednesday". In: *The Philadelphia Inquirer*. URL: https://www.inquirer.com/

news/food-stamps-hunger-federal-government-shutdown-20190116.html (visited on 03/03/2020).

Makioka, Ryo (2018). "Essays on Applied Microeconomics". PhD thesis. The Pennsylvania State University.

Pennsylvania Department of Human Services. *SNAPFacts*. URL: https://www.dhs.pa.gov/Services/Assistance/Pages/SNAP.aspx (visited on 03/03/2020).

Price, Donald I. and E. Shawn Novak (2000). "The Income Redistribution Effects of Texas State Lottery Games". In: *Public Finance Review* 28.1, pp. 82–92.

Rosenbaum, Dorothy and Brynne Keith-Jennings (Oct. 24, 2013). "November 1 SNAP CutsWill Affect Millions of Children, Seniors, and People With Disabilities: State-by-State Figures Highlight the Impacts Across the Country". In: URL: https://www.cbpp.org/sites/default/files/atoms/files/10-24-13fa.pdf (visited on 03/04/2020).

USDA (Jan. 2019). *USDA Announces Plan to Protect SNAP Participants' Access to SNAP in February*. URL: https://www.usda.gov/media/press-releases/2019/01/08/usda-announces-plan-protect-snap-participants-access-snap-february.

USDA (2020). "Supplemental Nutrition Assistance Program (SNAP) Monthly Issuance Schedule for All States and Territories". In: URL: https://fns-prod.azureedge.net/sites/default/files/snap/Monthly-Issuance-Schedule-All-States.pdf (visited on 03/03/2020).

3.A Data Appendix

SNAP Store Eligibility Data

SNAP benefits cannot be redeemed at all stores, but eligibility requirements are relatively relaxed. Many gas stations and convenience stores are eligible to accept SNAP benefits. Stores must be "authorized" by the FNS in order to accept SNAP benefits. In order to be eligible, stores must meet at least one of two staple food requirements, which are defined below.[71,72] Staple foods are defined as "the basic foods that make up a significant portion of a person's diet and are usually prepared at home and eaten as a meal." The 4 categories are (1) fruits or vegetables; (2) meat, poultry, or fish; (3) dairy products; and (4) breads or cereals. "Perishable" foods are defined as foods that "would spoil or suffer significant deterioration in quality within 2-3 weeks at room temperature." Shelf-stable foods that would deteriorate until after being opened are not counted as "perishable."

Criterion A (staple food inventory): Stores must stock 3 "stocking units" (e.g., cans, bags, bunches, etc.) of 3 "varieties" of each of the 4 staple food categories. This requires that they have 3 "stocking units" of 1 perishable food in 2 staple food categories. In aggregate, this requires the store to have 36 stocking units (6 of which must be perishable) of staple foods. A minimal assortment that would satisfy these requirements would be the following (this has 1 more perishable variety than necessary):

- **Fruits/Vegetables**: 3 cans of tomatoes, 3 cans of pineapple, 3 packages of frozen tater tots (perishable)
- **Meat**: 3 cans of tuna, 3 packages of beef jerky, 3 packages of frozen chicken nuggets (perishable)
- **Dairy**: 3 packages of powdered milk, 3 canisters of grated Parmesan, 3 cartons of milk (perishable)

[71] These requirements were adopted on December 15, 2016 and made changes to the Criterion A and B requirements. However, Sec. 765 of the Consolidated Appropriations Act of 2017 prevented some of these rules from going into effect and reverted requirements to prior to the 2014 Farm Bill. See RPMD Policy Memorandum 2018-04 for details. There may be some useful variation in this policy change as well.

[72] Some stores that do not meet these criteria can still be eligible if they are in areas with "significantly limited access to food." The store must not be a restaurant and meet other SNAP eligibility requirements. Stage 1 of this selection process requires the firm to be in or immediately adjacent to a low-income, low-access census tract, which is defined by the USDA Economic Research Service's Food Access Research Atlas. Upon passing Stage 1, the firm goes to Stage 2 which computes a Individualized Need for Access Calculator Tool that scores a firm based on proximity to meeting Criteria A and B, distance from other SNAP authorized stores, vehicle access rate per the American Community Service, firm's hours of operation, and history of SNAP violations by the firm owner or at the firm's location. Upon exceeding the INFACT threshold, a firm will be granted SNAP eligibility of a year, after which it must be re-evaluated. See RPMD Policy Memorandum 2018-03 for details.

- **Bread/Cereals**: 3 boxes of pasta, 3 boxes of cereal, 3 bags of rice

Criterion B (staple food sales): Staple food sales must make up more than 50% of gross retail sales. There are no stocking requirements, stores must simply meet the sales requirement.