Saving Space: Bulk Buying, Storage Costs, and Inequality

E. Mallick Hossain

May 28, 2019

*

1 Introduction

Buying in bulk is an effective way to save money, especially for storable items and they are found on almost all common retail product categories.

Despite these savings, I find that households making less than $25k are 6 percentage points less likely to make bulk purchases across a wide range of product categories, with especially large gaps in categories like toilet paper (20 pp) and paper towels (20 pp). For storable items with relatively inelastic demand, this presents a puzzle: Why are households not taking advantage of bulk discounts to reduce their expenditures?

*I wish to thank my wonderful advisers, Aviv Nevo, Holger Sieg, Katja Seim, and Sarah Moshary for their sound advice, constructive criticism, and immense patience throughout this writing process. They are prized mentors. I would also like to thank Emek Basker, Chris Cronin, Frank DiTraglia, Eileen Divringi, Jim Ferry, Ishan Ghosh, Joao Granja, Jeff Lin, Paolo Martellini, Davin Reed, Paul Sangrey, Anna Tranfaglia, and Keith Wardrip for their comments, questions, and extensive discussions on this topic which have helped to significantly improve this paper. I would also like to thank the Community Development and Regional Outreach group at the Philadelphia Fed for their comments and suggestions. Researcher(s) own analyses calculated (or 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. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

This paper identifies various factors and their importance in generating inequality in shopping for basic household necessities. Most importantly, it provides the detailed overview necessary to understand consumption inequality. Most research has focused on differences in aggregate spending between households or between various product categories. Coibion, Gorodnichenko, and Koustas (2017) show how these aggregates may mechanically overstate consumption inequality depending on how frequently a household's expenditures are measured. Pistaferri (2015) notes that "real" consumption inequality may still not be properly measured because a household's consumption may differ markedly from their expenditures. Aguiar and Hurst (2007) supports this assertion by showing that home production and time support a household's consumption even if measured expenditures fall during retirement.

As Aguiar and Hurst (2007) highlights, households can reduce expenditures without reducing their actual consumption. Argente and Lee (2017) shows that high-income households can better smooth their expenditures by substituting to cheaper brands. Low-income households are less able to do this because there are fewer cheaper brands to substitute down to. This paper builds on these threads to identify other common ways households can do that and how relying on aggregate expenditure measures can understate inequality.

I find that low-income households are "leaving money on the table" compared to their richer counterparts. Most notably, low-income households could reduce their expenditures without changing their consumption basket if they bought larger quantities of items. In my modeling section, I identify contributors that limit their ability to take advantage of these discounts.

2 Data

2.1 Nielsen Consumer Panel Data

I use the Nielsen Consumer Panel Dataset from 2004–2017 available through the Kilts Center for Marketing at the University of Chicago Booth School of Business. This dataset is a longitudinal panel of about 174,000 unique households, of which we observe about 40,000 each year from 2004–2006 and about 60,000 households each year from 2007–2017. This data records all purchases from any outlet intended for personal, in-home use and covers about 30% of consumer expenditures. About 1.5 million unique items (defined by UPC code) are recorded in categories such as groceries, cleaning supplies, health/personal care items, and basic general merchandise. The panel is based on a stratified, proportionate sample designed to be projectable to the United States population. It is balanced on demographic characteristics including household size, income, education, children, race, and occupation.

Households scan all items that they purchase, input quantities and prices (if necessary), date of purchase, and store purchased from. Households are incentivized to stay active in the panel by monthly prize drawings, points for data transmission, and sweepstakes as well as ongoing communication from Nielsen to ensure cooperation and address any problems. Nielsen retains about 80% of its panel from year to year with the mean and median tenure of a household being 4 and 3 years, respectively. To further ensure data quality, Nielsen institutes a minimum purchase threshold based on household size that must be met to be deemed "active".

For my analysis, I exclude households with a student or military head of household as well as any purchases of alcohol, "deferred" modules which Nielsen has stopped tracking, or "magnet" products.[1] Furthermore, I drop any modules for which fewer than 3 unique sizes are purchased or fewer than 100 purchases are recorded in a given year. This leaves me with 957 product modules.[2] My descriptive statistics and reduced-form analysis are generated from this sample.

[1]Magnet products are typically non-packaged, variable weight products like fresh produce or items ordered from a store's deli counter.

[2]A product module is a narrow product category defined by Nielsen. For example, toilet paper, paper towels, and eggs are separate product modules.

3 Stylized Facts

I document 3 stylized facts to motivate my analysis:

1. 99% of product categories with multiple available sizes have quantity discounts and storable items have larger discounts.
2. Households making over \$100k are 6 percentage points more likely to buy in bulk compared to households making under \$25k. Gaps are particularly large for essential storable items like toilet paper, diapers, and paper towels.
3. Households increase their bulk buying when warehouse clubs open nearby, but this only applies to middle- and high-income households.
4. High-income households use bulk discounts to offset price increases from quality improvement and store choice.

3.1 Bulk Discount Prevalence

In order to compute the prevalence and magnitude of bulk discounts, I analyze purchases recorded in the Nielsen Consumer Panel data from 2004-2017. First, I compute the unit price for each item purchased. If multiple packages were purchased, I divide the total price paid by the number of packages to get the per-package price, which is then used to compute the unit price.[3] Hence, all products within a product module are mapped into their corresponding unit price (e.g. cents per ounce for milk, cents per roll for toilet paper).

Given the unit price, I then compute the quantity discount by estimating the following regression for each product module:

$$log(unitPrice_{ismt}) = \beta_1^m log(packageSize)_{iht} + \lambda_s + \lambda_m + \lambda_t + \epsilon_{ist}, \qquad (1)$$

[3] Purchasing multiple packages generates linear pricing and without this correction would likely understate the magnitude of bulk discounts. To the extent that multiple packages were purchased because of a promotion (e.g. buy 2 get 1 free), I map that promotion into its equivalent per-package price.

where *unitPrice* is the unit price of product i purchased on shopping trip s in market m at time t. *packageSize* is the amount available in a particular package (measured in the most common units of that product module). λ are fixed effects for brand-retailer (a shopping trip), market, and year-month. A market is a Nielsen Designated Market Area which is a non-overlapping county grouping that covers the United States.[4]

Figure 1 shows the distribution of the bulk discounts, as captured by β^m.

[Figure 1 about here.]

Based on this estimation, about 99% of product categories have evidence of bulk discounts with the largest on products like jellies, jams, and canned fruit and the smallest on categories like tomato sauce, tomato paste, and yogurt. By construction, this estimate excludes any products where less than 3 unique sizes were purchased.[5] The peak is primarily composed of non-storable products while the fat left tail is composed of storable products. Without storage costs or other frictions, these discounts on storable products give households an easy way to lower costs without having to sacrifice consumption.

3.2 Discounting Behavior

In addition to documenting the prevalence of bulk discounting, I then document how takeup of these savings options correlates with income. For each product purchased, I add an indicator of whether that product was purchased with a coupon, purchased on sale, purchased in a bulk size, or was a generic brand (a product can have more than one indicator). A "bulk" size is defined as a product in the top quartile of the size distribution for that product category. For each household, I take the expenditure-weighted average of each discounting behavior.

[4] A map of these DMA regions is available at https://www.thevab.com/wp-content/uploads/2017/06/2016-2017TVDMARegionMap_Small_v4.pdf.

[5] One might be concerned about the manufacturer's decision of what sizes to produce and in some cases, the manufacturer could have strategically chosen not to produce multiple sizes. This would bias the prevalence of bulk discounts upward. Including all non-deferred, non-alcohol modules reduces this number to 86%, which is still a vast majority of product categories.

Using this household-level discount propensity, I estimate the following pooled regression (essentially generating a binned scatterplot):

$$Y_i = \beta_1 Income_i + \beta_2 Age_i + \beta_3 HHSize_i + \beta_4 Child_i + \epsilon_i, \tag{2}$$

where $Income$ is a dummy for a household i's income bin, Age is the head of household's age, $HHSize$ is the number of people in the household, and $Child$ is an indicator for whether the household has any children.[6] The results are plotted in Figure 2 and they show that overall, richer households are more likely to take advantage of bulk purchases, and sales, but are less likely to purchase generic brands. The fact that generic brands and bulk purchases are more prevalent than sales is supported by Griffith et al. (2009). Additionally, these differences are primarily driven by storable items.[7] The gap for storable items is 7 percentage points compared to 3 percentage points for non-storable items. Buying in bulk is the most commonly used discount option and is often available even when coupons and sales are not available. Some of this correlation could be driven by the possibility that richer households are just consuming more of these products and hence choose to buy larger sizes because of this increased consumption. However, for products where this is unlikely, such as toilet paper, paper towels, and diapers, these patterns persist, suggesting low-income households are paying more for the same basket simply because they are buying smaller quantities per trip.

Across popular spending categories, the biggest gaps are in categories like toilet paper (20 pp), paper towels (20 pp), and diapers (13 pp). These are much larger gaps than in common non-storable items like milk (2 pp) and eggs (0 pp). Any gaps in non-storable items could be attributed to different consumption rates not captured by age, household size, or presence of children, because these items will perish if not consumed within a short period of time. On the

[6] I control for these covariates because household income varies with household size, age, and children for reasons outside of socio-economic status.

[7] I define storable items as those that can remain unopened at room temperature for at least 2 months without significant deterioration. Milk, eggs, and frozen items are not storable while soft drinks, cereal, detergent, and toilet paper are storable.

other hand, storable items do not have to be consumed immediately. While I cannot rule out consumption differences for all items, it would be unlikely that consumption rates would differ so greatly for these items. In a later section, I show that income is not predictive of toilet paper consumption after controlling for age, household size, and presence of children.

[Figure 2 about here.]

3.3 Discount Savings

The previous section established that high-income households are more likely to take advantage of various discounting strategies, except for buying generic brands. In this section, I establish that low-income households are leaving money on the table as a result of not taking advantage of these money-saving strategies and that these savings are more pronounced for storable products. In order to quantify the savings, I estimate the following regression (modeled after Griffith et al. (2009)):

$$\begin{aligned} log(UnitPrice)_{ismt} =& \beta_0 + \beta_1 Coupon_{ismt} + \beta_2 Sale_{ismt} + \beta_3 Generic_{ismt} + \quad (3) \\ & \sum_{q=2}^{4} \beta_4^q Quartile_{ismt}^q + \lambda_s + \lambda_m + \lambda_t + \epsilon_{ismt}, \quad (4) \end{aligned}$$

where $log(UnitPrice)$ is the log unit price of product i purchased on shopping trip s in market m at time t. $Coupon$, $Sale$, and $Generic$ are all indicators for whether that item was purchased with a coupon, on sale, or was a generic brand, respectively. $Quartile^q$ is an indicator for whether the item's size is in the second, third, or fourth quartile of the size distribution of that product module. λ are shopping trip (retailer-brand), market, and year fixed effects.

This specification gives the average discount relative to the average unit price of a particular brand (e.g. average per-sheet price of Charmin toilet paper at Target).[8] The coefficient estimates are reported below in Figure 3. Savings from

[8] I also estimate different specifications with retailer-category or retailer-UPC fixed effects and the results are similar.

each of these modes is substantial with coupons, generics, and bulk purchases saving the most money. The fact that generic brands and bulk purchases also provide the most savings is supported by Griffith et al. (2009). However, bulk discounts can offer even more savings and storable items have much larger savings than non-storable items.

[Figure 3 about here.]

4 Reduced-Form Analysis

4.1 Warehouse Club Entry on Bulk Purchasing

One possible explanation for this “bulk buying gap” is that low-income households may not have access to large, bulk sizes that high-income households do. Warehouse club stores offer access to large packages with low unit prices in exchange for an annual membership fee. The largest warehouse club chains are Costco, Sam’s Club, and BJ’s. If access is the primary driver of this gap, then warehouse club entry should correspond with an increase in bulk purchasing.

Over my sample period, I observe 419 warehouse club openings. While there are some endogeneity concerns about what drives the decision to open a warehouse, I leverage the panel structure of the data to avoid this issue because it is unlikely that an opening would be planned exactly when a household changes its consumption patterns. I use an event study framework to look at how bulk purchasing changes within a household before and after a warehouse club opens. I estimate the following regression:

$$Y_{imt} = \beta S_{it} + \gamma X_{it} + \lambda_{im} + \lambda_{mt} + \epsilon_{imt}, \tag{5}$$

where Y is the expenditure-weighted share of bulk purchases made by houeshold i in market m in quarter t. S is the number of warehouse clubs that have entered within 30 kilometers of the household as of month t.[9] X is a vector of

[9] Distance is measured between the centroids of the household ZIP code and the store ZIP code.

household characteristics including income, household size, age, and presence of children. λ are fixed effects for household-markets (to ensure identification is from club entry and not household moves) and market-quarter. Standard errors are clustered at the household-market level.

Before estimating Equation 5, I graphically illustrate the event study by plotting the coefficients of a modified version of Equation 5.

$$Y_{imt} = B_{it}\left(\sum_{t} \beta_t E_{it} + \alpha\right) + \gamma X_{it} + \lambda_{im} + \lambda_{mt} + \epsilon_{imt}, \tag{6}$$

where B is an indicator for whether a household was continuously located in a market with a warehouse club opening and was there during the 4 quarters prior to the opening and 8 quarters after the opening. E is a dummy for each of the 4 quarters prior to opening and each of the 8 quarters after opening. The β coefficients are plotting in Figure 4

[Figure 4 about here.]

The event study shows that warehouse clubs increase a household's bulk purchasing, but this effect is primarily concentrated in higher-income households, with households making under \$25k having no significant change in their purchasing behavior. Households making over \$100k exhibit the largest increase in bulk purchasing, but we cannot definitively say if their change is larger than the change for other households making over \$25k.

The estimation results for Equation 5 are reported in Table 1 and after a warehouse club opens, bulk purchasing increases by about 1 percentage point overall, but when this analysis is run separately on each quartile, we see that this effect is primarily on the upper end of the income distribution with households making under \$25k exhibiting no significant difference in their bulk purchasing.

One reason why low-income households may not change their bulk purchasing is that they cannot afford the warehouse club membership fee and therefore still do not have access to the club even if it is nearby. Even if that was not the case, the coefficient estimates on wealthier households suggest that increased access to bulk sizes is unlikely to substantially close the 6 percentage point gap in bulk purchasing behavior.

[Table 1 about here.]

4.2 Price Decomposition

How important are these factors in generating savings? How much money are households "leaving on the table"? If products were homogenous, one could simply examine differences in the unit price paid by high- and low-income households. Increased uptake of bulk discounts would be reflected in lower unit prices paid by high-income households. However, in reality, these savings can easily be masked, or even offset by substitution to higher priced brands or stores. In fact, this is exactly what happens in reality with high-income households more likely to purchase higher quality items or shop at premium stores. Therefore, in order to better understand how each of these factors contributes to the prices paid by households, I decompose the unit price of each product module into store, brand, and size components using a hedonic regression. The contribution of each component is then aggregated up to the household level based on that module's basket share and these are aggregated across households using Nielsen's projection weights. The hedonic regression is shown in Equation 7.

$$unitPrice^{m}_{brq} = \lambda_b + \lambda_r + \lambda_q + epsilon_{brq} \tag{7}$$

The hedonic regression decomposes the unit price of a product purchased in module m of brand b at retailer r in size quartile q into brand, retailer, and size components. All coefficients are relative to the reference category which is the smallest generic brand product sold at the most popular discount retailer. Figure **??** shows that the contribution of brand and retailer components increase the price for high-income households, but that size components offset a large portion of this increase. Most importantly, the figure illustrates the large importance of size components in determining an item's unit costs, dramatically overshadowing savings that come from sales or coupons.

5 References

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

Argente, David, and Munseob Lee. 2017. "Cost of Living Inequality During the Great Recession."

Coibion, Olivier, Yuriy Gorodnichenko, and Dmitri Koustas. 2017. "Consumption Inequality and the Frequency of Purchases." Working Paper 23357. Working Paper Series. National Bureau of Economic Research. http://www.nber.org/papers/w23357.

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

Pistaferri, Luigi. 2015. "Household Consumption: Research Questions, Measurement Issues, and Data Collection Strategies." *Journal of Economic and Social Measurement* 40 (1-4). IOS Press: 123–49.

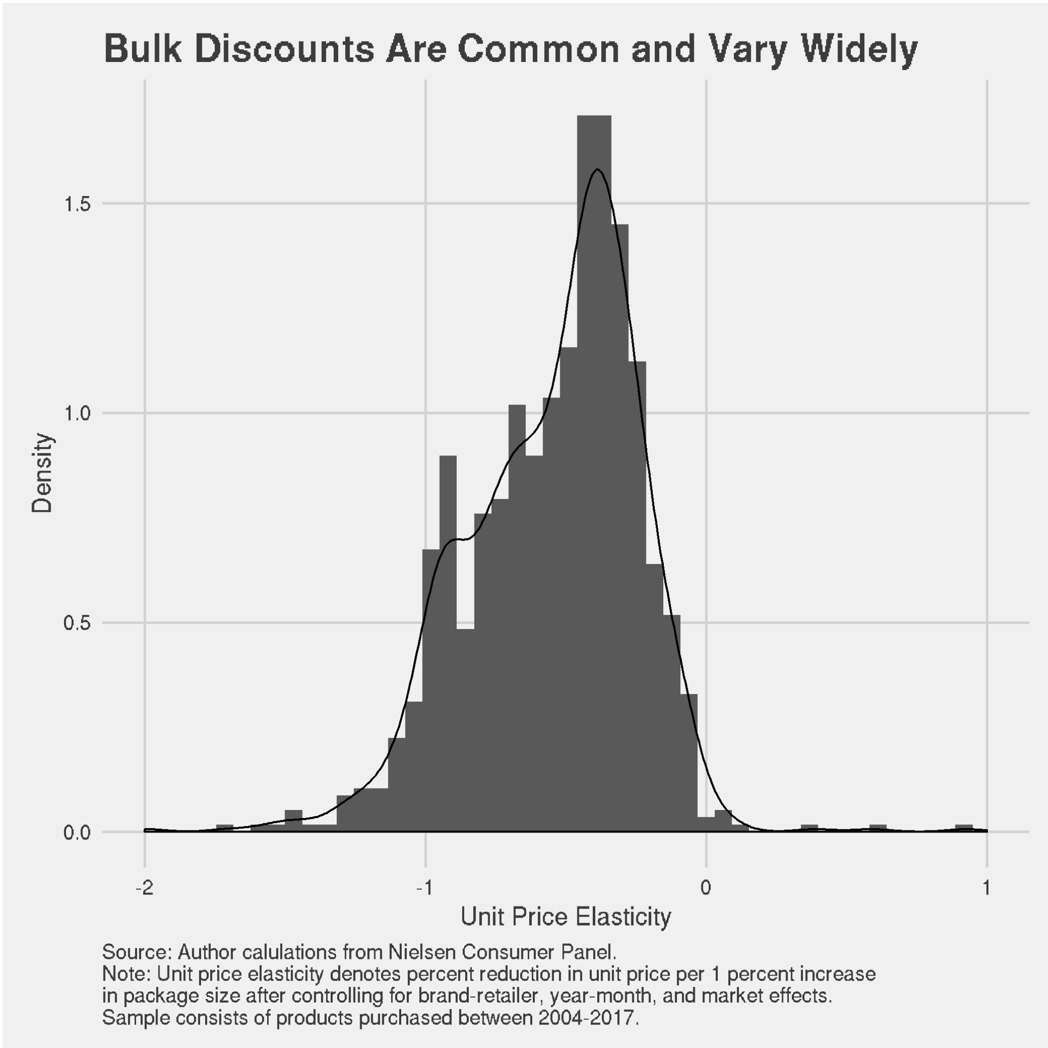
Bulk Discounts Are Common and Vary Widely
1.5
1.0
0.5
0.0
Density
-2
-1
0
1
Unit Price Elasticity
Source: Author calulations from Nielsen Consumer Panel.
Note: Unit price elasticity denotes percent reduction in unit price per 1 percent increase
in package size after controlling for brand-retailer, year-month, and market effects.
Sample consists of products purchased between 2004-2017.

Figure 1

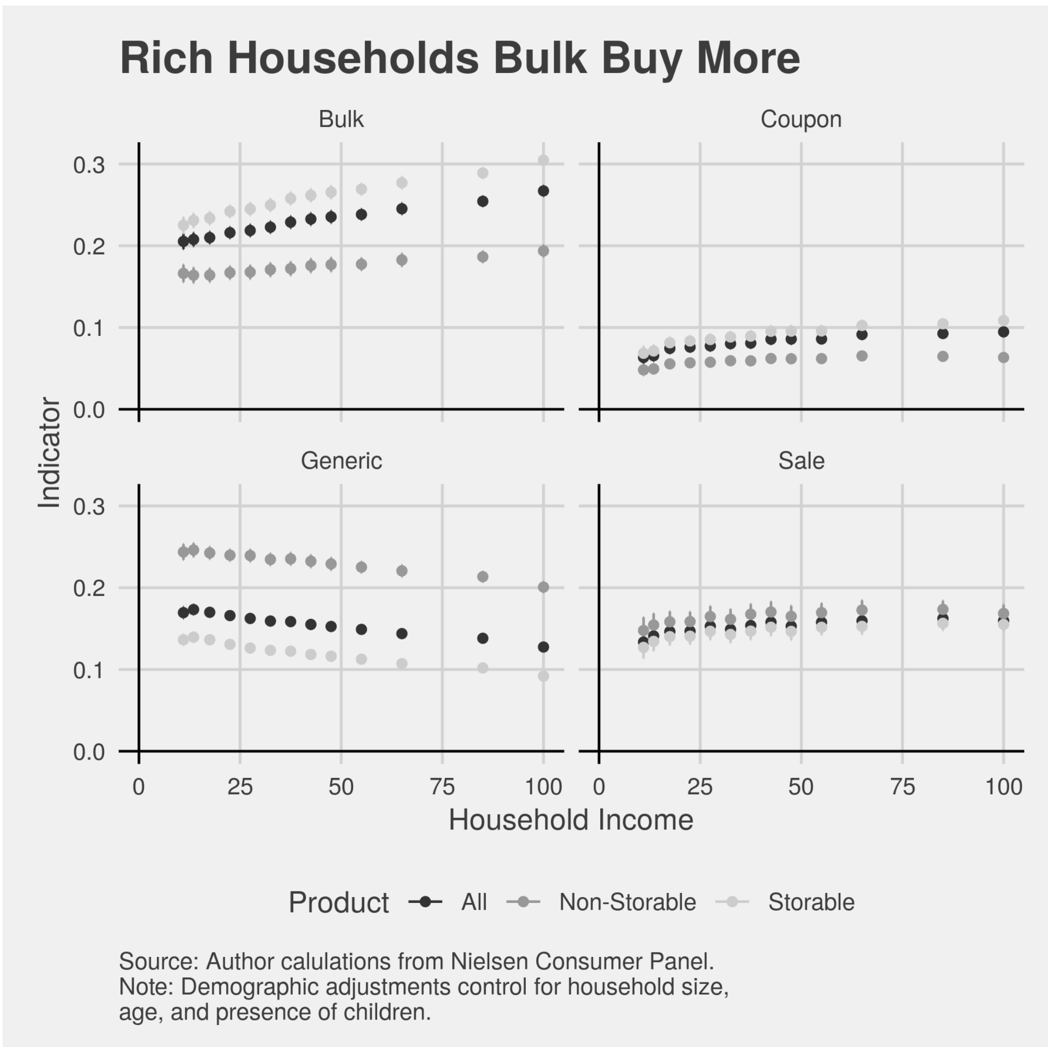
Rich Households Bulk Buy More
Bulk
Coupon
Generic
Sale
Indicator
0.0
0.1
0.2
0.3
0
25
50
75
100
Household Income
Product
All
Non-Storable
Storable
Source: Author calulations from Nielsen Consumer Panel.
Note: Demographic adjustments control for household size,
age, and presence of children.

Figure 2

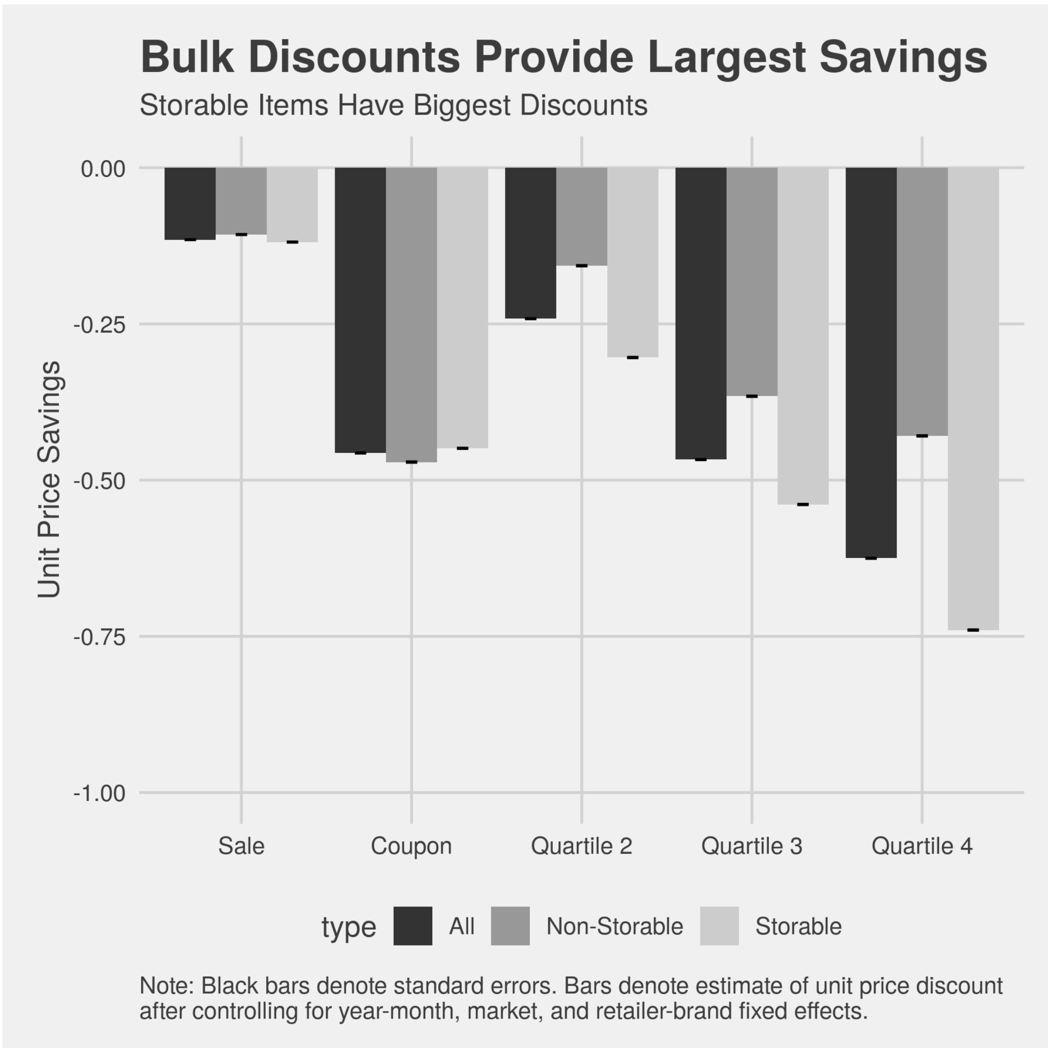
Bulk Discounts Provide Largest Savings
Storable Items Have Biggest Discounts
Unit Price Savings
0.00
-0.25
-0.50
-0.75
-1.00
Sale
Coupon
Quartile 2
Quartile 3
Quartile 4
type
All
Non-Storable
Storable
Note: Black bars denote standard errors. Bars denote estimate of unit price discount after controlling for year-month, market, and retailer-brand fixed effects.

Figure 3

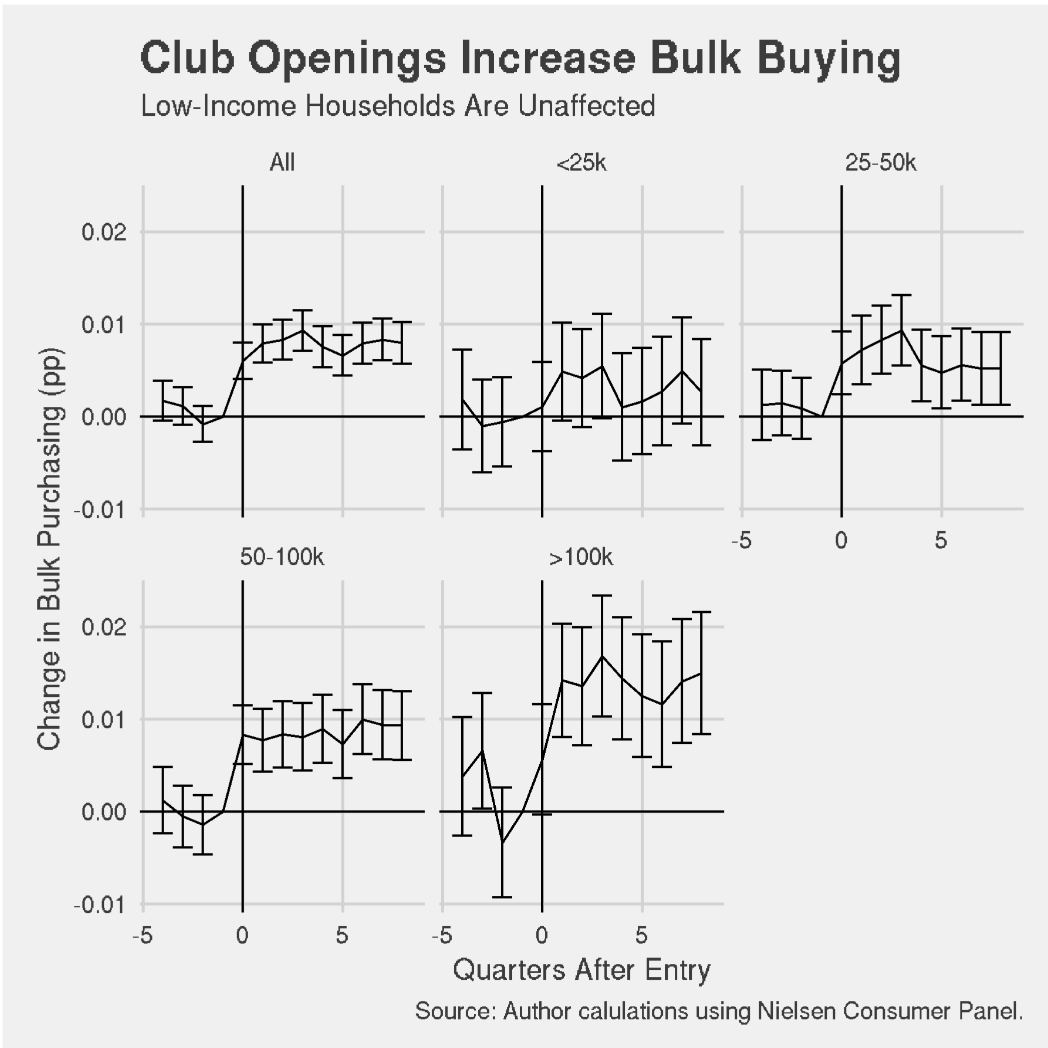
Club Openings Increase Bulk Buying
Low-Income Households Are Unaffected
All
<25k
25-50k
50-100k
>100k
Change in Bulk Purchasing (pp)
0.02
0.01
0.00
-0.01
-5
0
5
Quarters After Entry
Source: Author calulations using Nielsen Consumer Panel.

Figure 4

Table 1: Warehouse Club Opening

	All	<25k	25-50k	50-100k	>100k
	(1)	(2)	(3)	(4)	(5)
<20 km	0.010***	0.0003	0.009***	0.012***	0.013***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
20-40km	0.007***	−0.001	0.008***	0.008***	0.012**
	(0.002)	(0.003)	(0.003)	(0.003)	(0.005)
Household-Market FE	Y	Y	Y	Y	Y
Market-Quarter FE	Y	Y	Y	Y	Y
Observations	3,027,655	424,285	963,860	1,182,324	457,186
Adjusted R^2	0.673	0.642	0.671	0.689	0.694

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are clustered at the household-market level.