Saving Space: Bulk Buying, Storage Costs, and Inequality

E. Mallick Hossain

April 22, 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 paper towels (20 pp), toilet paper (20 pp), and diapers (13 pp). For storable items with relatively inelastic demand such as toilet paper or diapers, this presents a puzzle. Why are households not taking advantage of the bulk discounts available to reduce their expenditures?

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 or "deferred" modules which Nielsen has stopped tracking. 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.[1] My descriptive statistics and reduced-form analysis are generated from this sample.

3 Stylized Facts

I document XX 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 use bulk discounts in addition to taking advantage of sales and coupons

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

compared to households making under $25k. Gaps are particularly large for essential storable items like toilet paper, diapers, and paper towels.

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.[2] 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:

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

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.[3]

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

[2]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), maps that promotion into its equivalent per-package price.

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

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.[4] 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. 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.[5] 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

[4]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.

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

prevalent than sales is supported by Griffith et al. (2009). Additionally, these differences are primarily driven by storable items.[6] 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 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.]

–>

–> –>

–>

–> –>

[6] 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.

–>

–>

–>

–> –>

–> –> –> –> –> –>

–> –>

–>

–> –>

–> –> –>

–>

–>

–>

–>

–>

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.

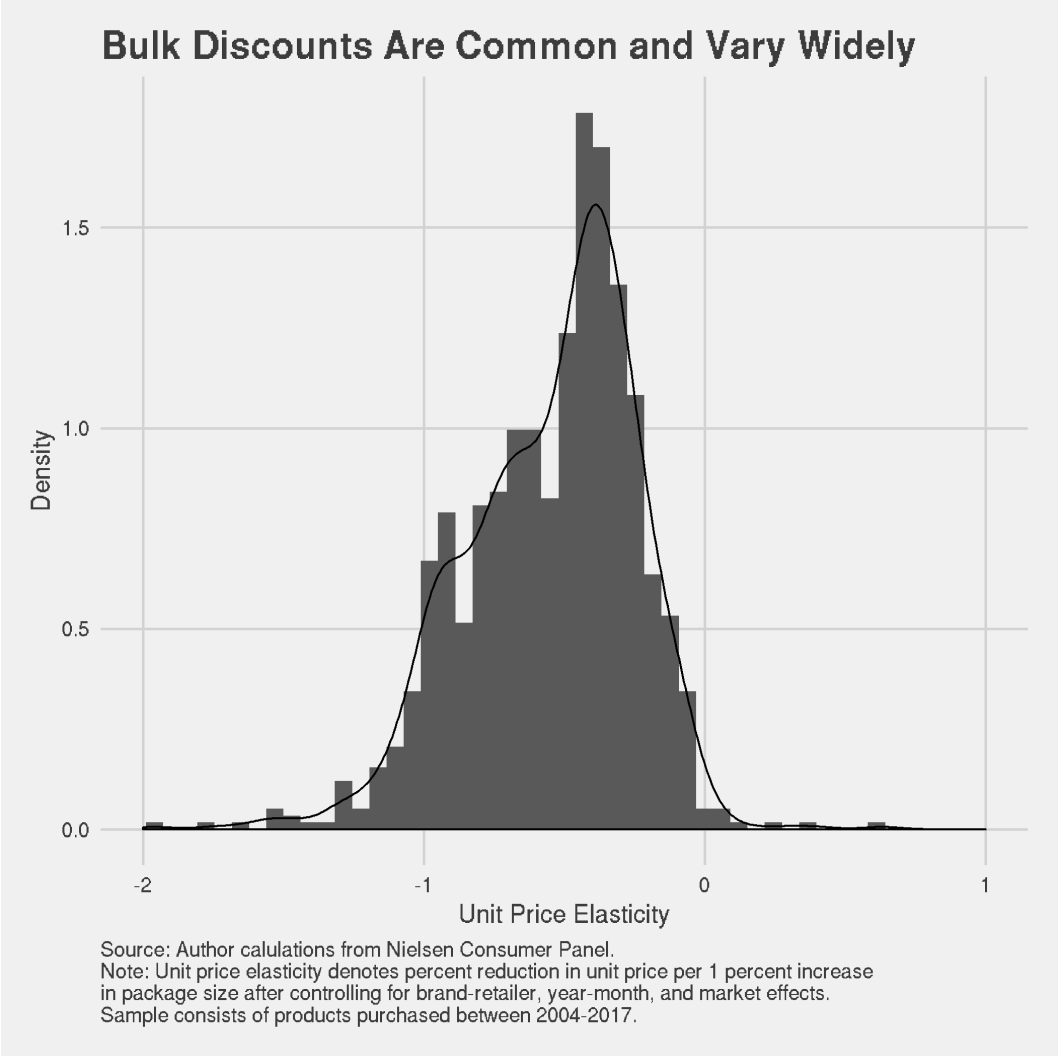
Bulk Discounts Are Common and Vary Widely
Density
1.5
1.0
0.5
0.0
-2
-1
0
1
Unit Price Elasticity
Source: Author calulations from Nielsen Consumer Panel.
Note: Unit price elasticity denotes percent reduction in unit price per 1 percent increase
in package size after controlling for brand-retailer, year-month,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

Figure 1

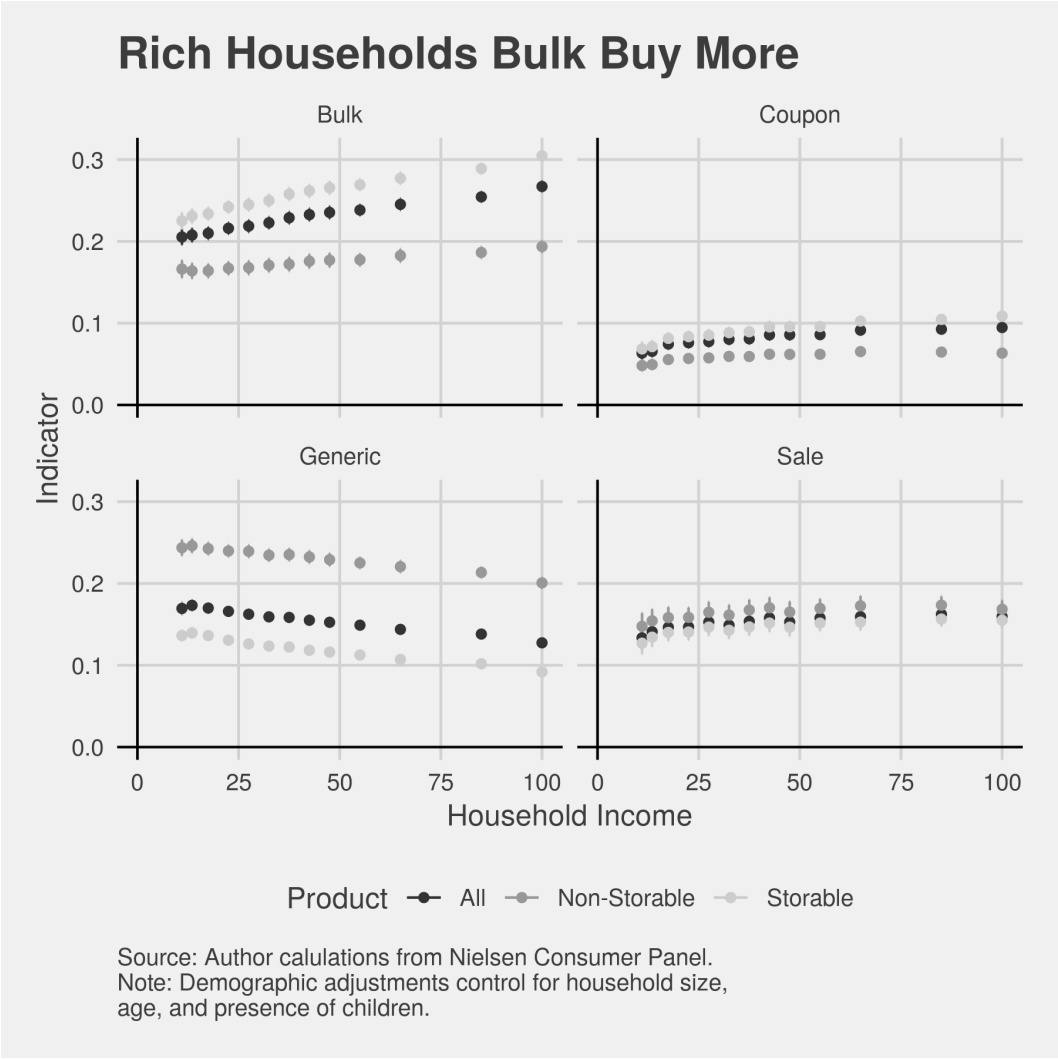
Rich Households Bulk Buy More
Bulk
Coupon
Generic
Sale
Indicator
0.0
0.1
0.2
0.3
0
25
50
75
100
Household Income
Product
All
Non-Storable
Storable
Source: Author calulations from Nielsen Consumer Panel.
Note: Demographic adjustments control for household size,
age, and presence of children.

Figure 2