Bulk Buying and Inequality

E. Mallick Hossain

05 March, 2019

Abstract

Buying in bulk is a common way to save money, especially for storable items like toilet paper. Using granular household shopping data, I find that households making over \$100k buy 30% larger packages than households making less than \$25k. Furthermore, the relationship between quantity purchased and income is monotonically increasing. I estimate that low-income households spend 7% more for the same annual quantity than they would if they made larger per-trip purchases. By highlighting the importance of budget constraints and storage costs, this research highlights key structural factors that contribute to the "poverty penalty". Given the prevalence of bulk discounts, this research suggests that quantity discounts contribute to consumption inequality and may affect how we measure prices faced by different socio-economic groups. This research implies that reducing budget constraints would help low-income households achieve lower spending for the same basket of goods because they could buy larger quantities at lower unit costs.

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, Ishan Ghosh, Joao Granja de Almieda, Paolo Martellini, and Paul Sangrey for their comments, questions, and extensive discussions on this topic which have helped to significantly improve this paper. 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.

1 Introduction

Households have a variety of available options when trying to save money on common household purchases like groceries and general merchandise. They can clip coupons, search for lower prices, wait for sales, buy cheaper brands, or buy in bulk. Each of these options requires a different trade-off to achieve savings. Search strategies (coupons, sales, etc.) require households to spend time in exchange for lower prices. Brand substitution requires households to possibly sacrifice quality for savings. Both of these strands of strategies have been extensively studied. However, consumer behavior in response to bulk discounts has received less attention, especially for storable items where actual consumption may not respond strongly to lower prices. The trade-off required for a bulk discount is less obvious than for sales or brand substitution. On one hand, any forward-looking, expenditure-minimizing households with reasonable expectations of their own demand would purchase the package that delivers the lowest unit cost (i.e. the largest package). However, given that we do not see consumers pool on the largest size, what frictions could generate this pattern? One possibility is budget constraints. A bulk discount requires an upfront investment in exchange for future savings. The other friction could be storage or transportation costs. Buying a large quantity requires having the means to transport and store a large quantity.

As mentioned above, given the relatively low price of common household goods, relative to other expenses, any rational agent would appropriately save and spend to achieve the lowest unit cost of these items. Low-income households would be especially sensitive to reaping any potential savings because the marginal dollar is much more valuable to them than to a higher-income counterpart. Therefore, economic intuition would identify storage and transportation costs as the driving factor for why some households cannot buy bulk quantities.[1]

[1] Informal conversations have suggested as much. With few exceptions, economists I have talked to inevitably identify storage costs as the main driver of why consumers do not buy in bulk while non-economists tend to identify budget constraints as the main driver. This

This paper find that high-income households are more likely to take advantage of bulk discounts compared to similar low-income households. For toilet paper purchases, households making over $100k purchase almost 30% larger packages compared to households making below $25k. As a result, these high-income households realize a 7% savings over their low-income counterparts simply as a result of buying in bulk.

Previous research has specifically analyzed how households substitute between brands and take advantage of sales (Argente and Lee 2017, Erdem, Imai, and Keane (2003), Hendel and Nevo (2006a), Hendel and Nevo (2006b), Hendel and Nevo (2013), Pesendorfer (2002)). Nevo and Wong (2018) explore how the propensity to take advantage of these savings changed during the Great Recession. However, most research assumes that unit costs are uniform for a particular product and households choose quantities given a fixed unit price. In reality, this is not true because unit costs differ with package size and larger packages offer lower unit costs. Aside from Griffith et al. (2009) and Nevo and Wong (2018), few papers have looked specifically at bulk discounts as a money-saving technique.

The closest paper to this one is Orhun and Palazzolo (2019) which documents differences in purchase behavior between high- and low-income households. By comparing spending patterns at the beginning and end of a month, they find that liquidity constraints are an important reason why low-income households cannot take advantage of sales. However, they also look at bulk purchasing, but they find that low-income households do not take advantage of bulk purchases even when they have more money at the beginning of a month. The precise cause of this gap remains elusive. Budget constraints could still be a factor, but the difference between liquidity at the start and the end of the month is still not enough to make the up-front investment necessary for a bulk purchase. The other explanation is that storage and transportation costs are the primary factors, so even with extra room in their budget, households have no place to keep a bulk quantity.

paper will help pin down the relative importance of these competing explanations.

This paper contributes to the literature by explicitly incorporating how a household's purchase decision may depend on both a product's price and its size and how these preferences may relate to demographic characteristics.

This paper documents three stylized facts:

1. Households making over $100k purchase 29% larger packages of toilet paper compared to households making less than $25k.
2. The relationship between household income and package size purchased is monotonically increasing in income.
3. Due to bulk discounts, households making over $100k save 7% annually compared to frequently buying small packages like low-income households do.

Overall, this paper quantifies the "poverty penalty" paid by poor households that results from them being unable to take advantage of the same discounts that rich households can. This is particularly stark when examining necessity goods for which total consumption is fixed and relatively predictable. Over the course of a year, rich and poor households purchase the same amount of toilet paper, but have significantly different expenditure paths, which ends up costing low-income households more. This paper provides evidence and analysis to explain these patterns.

2 Data

2.1 Nielsen Consumer Panel Data

I use the Nielsen Consumer Panel Dataset from 2006–2016 available through the Kilts Center for Marketing at the University of Chicago Booth School of Business. This data is a longitudinal panel of between 40,000-60,000 households which record all purchases from any outlet intended for personal, in-home use. Products can be characterized as household non-durables such as groceries, cleaning supplies, health/personal care items, and basic general merchandise.

The dataset is estimated to contain about 30% of consumer expenditures. About 1.5 million unique items (defined by UPC code) are present in the data. 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 are provided a scanner with which they 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 remove any households where the head of household is in the military to avoid instances where store choice may be limited (e.g. all shopping is on a military base) or students where shopping choices may be different than a typical household (e.g. a meal plan and dormitory housing with various amenities).

In order to examine effects related to the value of bulk purchasing that is separate from immediate consumption, this analysis focuses on a product that is storable, non-perishable, and universal in consumption, namely, toilet paper. While this research focuses on toilet paper consumption, its conclusions are generalizable to overall household consumption, toilet paper provides the cleanest product by which to conduct this analysis.[2] Based on CPI basket weights, storable, non-perishable items account for about 2-5% of total expenditures and about 10-23% of nondurable expenditures.[3] All of the analysis that follows is done on the subset of toilet paper purchases in my Nielsen data. My sample

[2]That is, before it is used. Pun intended.
[3]See Appendix for items and basket shares.

includes about 3.7 million observations across 120,000 households between 2006 and 2016.

Data cleaning and organization follow Orhun and Palazzolo (2019) and take efforts to identify and remove missing purchase information and incorrectly recorded values, which help address concerns raised in Einav, Leibtag, and Nevo (2010). Further details are in the Appendix.

Of the many brands of toilet paper sold each year, I focus on the top 5 "brands" (Angel Soft, Charmin, Cottonelle, Quilted Northern, and Scott 1000) which account for over 60% of purchases.[4] Private-label products account for about 25% of sales with the remainder composed by a long tail of other brands. The average prices and bulk discounts for the top 5 brands are in Table 1:

[Table 1 about here.]

Overall, we see that bulk discounts can be substantial, but require a much larger investment. Increasing purchase size from a 4 to a 6-roll package does not net much savings, but further increasing to a 12-roll package does generate substantial unit cost savings.

Table 2 shows the positive correlation in the raw data between income and package size purchased. As will be shown in the following section, this correlation is robust to controlling for a wide range of household characteristics as well as brand and store preferences.

[Table 2 about here.]

2.2 Nielsen Retail Scanner Data

I also use Nielsen's Retail Scanner Data in order to estimate a discrete choice model. The Nielsen Retail Scanner Data contains weekly pricing, volume, and promotional activity of about 35,000 stores from about 90 retail chains and

[4] A "brand" is defined by the name borne on the product. This is a slightly finer gradation than might be commonly expected in that Charmin and Charmin Essentials are different "brands" as are Scott 1000 and Scott Extra Soft.

covers 2006-2016. Each store in the panel reports their weekly quantity sold as well as the weighted average unit price. For a subset of stores, they also report their promotional activity (i.e. if a product was featured or on display). This data is essential to match a consumer's purchase decision with the product assortment they faced when making their decision. I am able to match the Retail Scanner Data with the Consumer Panel data based on store identification numbers that correspond to an individual store location and the week of the purchase. Overall, I am able to match about 30% of purchases to stores. This is in line with Nielsen estimates of matching between the two datasets. This primarily results from the fact that Nielsen only has data-sharing agreements with about 35,000 stores in 90 retail chains and households may shop at stores that do not share their data with Nielsen.

Since the Retail Scanner Data only records information on products that have positive sales, I cannot between a product was available and did not sell or a product that was unavailable. To pin down a consumer's choice set, I assume that only products with positive sales are in their choice set. Additionally, there are cases where a household records a purchase, but a corresponding sale is not found in the Scanner data. In these instances, I append the "missing" purchase to the store's sales in order to create the consumer's choice set.

3 Descriptive Evidence

3.1 Average Daily Consumption

One important characteristic of toilet paper that I use in this analysis is that consumption is relatively fixed by household and that increased package size purchases by high-income households do not reflect increased consumption. Furthermore, unlike other goods, consumption cannot be inter-temporally shifted and there is little scope for home production to substitute for this product. This should be intuitive from a biological perspective in that characteristics like household size and the presence of children should be most predictive of consumption, but not income.

In order to verify this, I first compute average daily consumption by computing the total volume of toilet paper purchased by a household, excluding the final purchase, and divide that by the number of days they were active in the panel. I then regress this household-specific consumption rate on household demographics to identify the most significant predictors. As an alternative validation, I use machine-learning (namely, 20-fold cross-validated LASSO) to identify the most useful predictors of consumption. LASSO is a method that identifies the most predictive subset of covariates in a regression and shrinks other coefficients to 0. Both analyses are detailed in the Appendix.

3.2 Stylized Facts

I document 2 stylized facts to motivate my analysis:

1. Households making over \$100k purchase 29% larger packages than households making under \$25k.
2. Quantity discount savings are significant and substantial, a 10% increase in package size reduces the unit price by 2%.

Combining these facts together, I calculate that low-income households spend 6% more over the course of a year because they do not take advantage of bulk discounts at the same rate as high-income households even after accounting for their consumption rate.

3.3 Package Sizes

I find that high-income households purchase larger package sizes than low-income households by estimating the following regression:

$$Y_{ikrbmt} = \beta Inc_{imt} + \gamma X_{imt} + \lambda_r + \lambda_b + \lambda_m + \lambda_t + \epsilon_{imt}, \tag{1}$$

where Y is the log standardized package size purchased made by household i on shopping trip k at retailer r of brand b in market m at time t.[5] Inc is the household income quantile and X is a vector of other household characteristics including household size, type of residence, marital status, race, ethnicity, age, and education. λ are fixed effects.

Column 1 of Table 3 shows that high-income households (those making over $100k) purchase packages that are about 27% larger than those purchased by the lowest-income households (those making under $25k), even after accounting for household demographics. This coefficient remains stable even after controlling for a household's consumption rate (Column 2). After accounting for store choice, the income coefficient decreases substantially (Column 2), but is still substantial and significant at 11%. Furthermore, package size is still increasing in household income. Column 3 adds brand fixed effects, and the remains stable. As households become richer, they are more likely to buy larger packages. Somewhat surprisingly, Figure 1 shows that even at finer income gradations, this pattern is monotonically increasing. Overall, these patterns suggest that there are multiple contributing factors to the decision to buy in bulk. Store choice is clearly an important factor, but I abstract away from that in this paper because there are many other factors such as shopping needs and geography that also affect this decision. I focus on the more puzzing within-store, within-brand purchase decision that is directly correlated to income. Put more succinctly, it is surprising that even for households within the same store, choosing the same brand, there is a package size difference that is correlated with income.

[Table 3 about here.]

[Figure 1 about here.]

[5] Packages are standardized to be in units of 225 sheet 2-ply rolls. Because private-label products have the same brand code in the Nielsen data, I recode the private label brand to correspond to a brand-retailer to better control for potential quality differences. As an illustration, this ensures that Great Value (Walmart brand) is distinct from Kirkland Signature (Costco brand).

3.4 Bulk Discounts

After establishing this stark difference in package size choice, I estimate the magnitude of bulk discounts using the following regression:

$$log(P)_{krbmt} = \beta log(Q)_{krbmt} + \lambda_r + \lambda_b + \lambda_m + \lambda_t + \epsilon_{krbmt}, \tag{2}$$

where P is the unit price paid on shopping trip k at retailer r of brand b in market m at time t.[6] λ are fixed effects.

Table 4 shows that overall, the bulk discount is such that a 1% increase in the package size decreases the unit cost by 0.25%. After accounting for brand and retailer differences, the bulk discount is lessened, but still substantial at 0.19%.

[Table 4 about here.]

One possible criticism is that this measure of the bulk discount is conditional on a household making a purchase and does not reflect the actual pricing behavior of stores. In order to ensure that this is a good approximation for actual pricing behavior, I run a similar regression on the Retail Scanner data. I modify the fixed effects to account for the store-level data I have and I include store, brand, and week fixed effects to isolate the bulk discounting available within a store and the results are strikingly similar to the bulk discounts reflected in the Consumer Panel. Both of these regressions reveal that a large amount of price variation is due to brand, with some remainder attributable to store or time.[7]

[Table 5 about here.]

[6] Packages are standardized to be in units of 225 sheet 2-ply rolls.

[7] Honestly, using almost half a billion observations to verify that a bulk discount exists and is meaningful is overkill. Anyone who has been grocery shopping could attest as much.

3.5 Estimated Savings

Combining together the differences in package sizes purchased and the estimated bulk discounts, we calculate that low-income households could reduce their unit costs and therefore their overall expenditures by about 6%.

I get a similar estimate from running the following regression as well, which provides suggestive evidence that high-income households are able to obtain similar standardized sizes at a lower unit cost.

$$log(P)_{ikrbmt} = \beta_1 Log(Q)_{krbmt} + \beta_2 Log(Q)_{krbmt} * Inc_{imt} + \gamma X_{imt} + \lambda_r + \lambda_b + \lambda_m + \lambda_t + \epsilon_{imt}, \quad (3)$$

where P is the log unit price paid by household i on shopping trip k at retailer r of brand b in market m at time t. Q is the number of standardized rolls purchased. Inc is the household income quantile and X is a vector of other household characteristics including household size, type of residence, marital status, race, ethnicity, age, and education. λ are fixed effects.

The β_2 coefficients are plotted in Figure 2. These coefficients show a steady negative relationship between household income and the unit price paid by a household starting once a household makes above $25k and decreasing to a 6% reduction in unit costs for the same size package after controlling for a wide range of both household demographics as well as various fixed effects.

[Figure 2 about here.]

4 Model

The above descriptive analysis suggests some underlying factor that is contributing to the positive relationship between income and package size. However, the mechanism or causal factor remains elusive. Even Orhun and Palazzolo

(2019) provides limited evidence that liquidity constraints are the cause of this difference. The most likely explanations rest on budget constraints or on storage/transportation costs. In order to assess the relative importance of these factors, I estimate a multinomial logit model that incorporates various product characteristics that could explain this difference.

4.1 Multinomial Logit

This is the workhorse model of any discrete choice setting. For simplicity and computational efficiency, I estimate this model on a restricted subset of the product space. In particular, I only consider the top 6 brands (Angel Soft, Charmin, Cottonelle, Quilted Northern, Scott, and private-labels) and 4 different sizes (4-, 6-, 12-, and 24-roll packages). This generates a total of 24 brand-sizes. While this may seem like a limited sample given the wide variety of brands and package sizes, this covers over 70% of consumer purchases that matched with the Retail Scanner data.

The setting is as follows. A household makes a shopping trip to a store in which they purchase toilet paper. In over 85% of shopping trips, households only purchase 1 package, so a discrete choice model is appropriate for this setting.[8] The household's choice set consists of all products with positive sales at the store they visit during the week of their visit. This is obtained by matching consumer panel data with retail scanner data at the store-week level.

I only model the product choice decision conditional on purchasing toilet paper. Since toilet paper is storable, one might want to also model the purchase/no-purchase decision allowing for an underlying inventory state. However, toilet paper tends to go on sale relatively infrequently and a majority of purchases are made at full price. This consideration would imply that my price elasticity

[8] Households purchasing 2 packages make up the majority of remaining purchases. This could be addressed in one of two ways. We could modify the choice set to include 2-package products. The other option is to treat my estimates as an upper bound on price sensitivity. Because of bulk discounts, purchasing multiple packages is weakly more expensive than the equivalent larger package. One other fact is that in 98% of trips, a single UPC is purchased.

estimates are a lower bound since the purchase/no-purchase decision would relax the constraint of having to make a purchase.

4.1.1 Utility

Households choose the product that maximizes their utility, which I represent as the following.

$$U_{njt} = X_{njt}\beta + \epsilon_{njt} \tag{4}$$

where U_{njt} represents the utility that household n gets from product j on trip t. X_{njt} is a matrix of product characteristics, namely price, ply, package size, and unit cost. ϵ_{njt} is an error term drawn from a Type I Extreme Value distribution. β reflects the population's average taste for each product characteristic. This model is estimated using MLE.

To streamline computation, I focus on Boston in 2014-2016, but hope to expand this sample more widely.

[Table 6 about here.]

[Table 7 about here.]

[Table 8 about here.]

The coefficients fit relatively well with intuition. The price coefficient is significantly negative as expected. Angel Soft is the reference brand, so the brand fixed effects indicate that only Scott is less likely to be bought, but all other brands, including private labels are more likely to be bought over Angel Soft. An increased package size slightly increases the likelihood of purchase while the days' supply offered by the package (which I define as the standardized package divided by household's average daily consumption) slightly reduces the probability of purchase.

To translate this into dollar terms, we can compute the willingness to pay by dividing each coefficient by the price coefficient. This implies that households are willing to pay an extra \$0.68 for a package with an extra unit, but they would have to be paid an extra \$0.11 to purchase an additional days' supply of toilet paper. Since days' supply is a function of package size, this implies and interesting trade off. For high consumption households, they prefer larger package sizes, but for low consumption households, they prefer smaller package sizes. The willingness to pay for particular brands does give large estimates which outstrip the usual price differences seen in a store (WTP ranges from +\$6.79 for Cottonelle to -\$3.84 for Scott).

Furthermore, while this model is a good starting point, a pure multinomial logit model implies a strict substition pattern as a result of the independence from irrelevant alternatives property. This implies that if a new product is introduced, whether a Charmin 36-pack or a private-label single roll, it will draw from other alternatives proportionately. This is at odds with what we would expect given that a Charmin 36-pack is more suited for customers that prefer Charmin or large packages, but it is unlikely to compete with, say, a Scott 4-pack.

4.2 Random Coefficients Model

As a baseline model, I estimate a discrete-choice mixed logit (random coefficients) model. The advantage of using a mixed logit model is that I can take advantage of the panel structure of my data and allow for individual heterogeneity in preferences across various product characteristics. For example, different households may have different preferences over price or package size. A standard discrete choice logit model restricts these preferences to be the same across all households. Furthermore, a mixed logit model also does not place restrictions on the kinds of substitution patterns that may be present.

Let U_{hj} represent the utility that household h gets from product j. We can decompose this utility into a "deterministic" part that relates the utility U to observable factors and an unobservable portion ϵ_{hj}.

$$U_{hj} = V_{hj} + \epsilon_{hj} \tag{5}$$

Letting V be represented as a linear function of observables X_{hj}, we get the following equation:

$$U_{hj} = \beta_h' X_{hj} + \epsilon_{hj} \tag{6}$$

where β is allowed to be different for each household h. For most of the individual elements of β, I will assume a normal distribution except for the price coefficient which I will assume to be lognormal in order to ensure that it is negative. The elements of x are product characteristics that will include ply, brand, package size, standardized unit cost, and total cost.

With this representation and assuming that ϵ follows a Type 1 extreme value distribution, we can write the choice probabilities as:

$$P_{hi} = \int \frac{\exp(\beta_h' x_{hi})}{\sum_j \exp(\beta_h' x_{hj})} f(\beta) d\beta \tag{7}$$

The key parameters of interest are the mean and variance of the distribution of each β. These estimates will illustrate the heterogeneity in preferences for various product features, with my focus being on preferences for package size and costs.

The next step is to relate this heterogeneity in individual-level preferences to demographic characteristics. Because of the panel structure of the data, we can obtain estimates of the individual-level preference parameters. Combining that with the richness of the demogrphic data will allow us to relate preference heterogeneity with demographics. For example, if there is substantial heterogeneity in package size preferences, but that preference is strongly correlated with housing, then this would suggest that these preferences are being driven by storage costs. Similarly, if preference heterogeneity in total price was strongly correlated with income, this would suggest preferences are being driven by budget considerations.

5 Counterfactuals

5.1 No Bulk Discounts

6 Robustness

6.1 Credit Access

If budget constraints are preventing low-income households from buying in bulk, then having access to a credit card may help them better smooth their expenditures over short time horizons. If budget constraints are binding over longer time periods, which could be due to seasonal employment, income volatility, or other unexpected income shocks, then credit cards would not provide the short-term liquidity because the interest rate would effectively cancel out potential bulk discount savings.

In order to test this hypothesis, I add an indicator for whether or not the household has a credit card. Most interestingly, the regression results in Table 9 suggest that credit access does not differentially affect households based on their income, but that having a credit card is associated with purchasing larger package sizes, regardless of income.[9] For the poorest households, this effect is large enough to overcome the difference in purchasing with high-income households.

[Table 9 about here.]

[9] This subsample is only available after 2013. As a further check, I rerun the original analysis on this subsample and the positive relationship between package size and income still holds.

6.2 Transportation Costs

Another factor that could be generating this pattern is transportation costs. Low-income households may more heavily rely upon public transport or have less access to cars than high-income households. In order to address this concern, I directly control for car ownership rates from the ACS. The regression results in Table 10 suggest that transportation may be an important factor in package sizes purchased, but the variation that is being captured by car ownership does not affect the income coefficients.

[Table 10 about here.]

6.3 Storage Costs

One competing explanation is that storage costs are driving differences between high- and low-income households. Equation 3 directly controls for the type of residence the household lives in (single-family home, apartment, or mobile home). However, there may be other storage aspects I am missing. High-income households likely have larger homes and therefore lower storage costs for large quantities of items like bulk toilet paper. In order to test this explanation, I redo the above analysis for tampons. Tampons are storable, non-substitutable, and demand is relatively predictable and fixed. However, storage costs are substantially lower for tampons because their packages are smaller. As an approximation, a 2-3 month supply of tampons (~50 units) takes up about 158 cubic inches of space while a 2-3 month supply of toilet paper takes up 5 times as much space.[10]

Estimating Equations 3 and 2 gives qualitatively similar results to my toilet paper analysis. Quantitatively, the results are smaller overall, with slightly flatter bulk discounts (-0.22 instead of -0.25) and less differences in package size purchases (10% difference instead of 24%). However, some of the package

[10] Volume calculations based on Playtex Sport 50-unit package and Cottonelle Ultra ComfortCare 12-unit package.

size differences could be due to the usage of pads. That analysis is not done here. The main takeaway is that the positive correlation between income and package size remains even when storage costs are substantially reduced.

For completeness, these size choices are likely not due to brand preferences because most brands offer a similar range of sizes (see Figure 2).

[Table 11 about here.]

6.4 Brand Preferences

One alternative explanation for this bulk purchasing pattern could be brand and quality preferences. Maybe cheap brands are offered in small sizes while expensive brands are larger. Therefore, what looks to be bulk purchasing could be brand preference. Figure 3 plots the package size distribution of the top 10 most popular brands. There is substantial overlap in the size distributions (with the exception of warehouse club private labels), so this pattern is unlikely to be driven by brand and quality preferences.[11]

[Figure 3 about here.]

6.5 Salience and Consumer Awareness

Following Chetty, Looney, and Kroft (2009), this could be a case of salience and consumer awareness in that consumers are not aware of the bulk discount because it is not presented to them or they do not compute unit prices when making purchases. To test this, I leverage state-level variation in laws about displaying unit-costs. 18 states have some form of law regulating the display of unit prices, 5 have guidelines, and 27 have neither guidelines or regulations. As my measure of unit-pricing laws, I use the information compiled in the National Institute of Standards and Technology (NIST) Handbook 130. This handbook is

[11] The top 10 brands account for over 80% of all purchases, so for clarity I plot only the distributions of these top 10 brands.

published annually following the National Conference on Weights and Measures and consists of numerous laws and regulations that are recommended for states to adopt in order to "achieve, to the maxiumum extent possible, uniformity in weights and measures laws and regulations among the various states and local jurisdictions in order to ... provide uniform and sufficient protection to all consumers in commercial weights and measures practices" (National Institute of Standards and Technology 2019). Every year, NIST publishes a summary of which states have adopted unit pricing laws. I use their classification to indicate whether a state has no laws, guidelines, or laws around unit pricing. Since 2004, very few states have changes their status, so the coefficients on unit pricing stringency is primarily being identified off of cross-sectional variation as opposed to longitudinal variation.

Table 12 provides evidence that laws and guidelines are associated with larger package size purchases. Whether the adopted policies are laws or guidelines does not appear to have a substantial difference, but there is a significant difference between having and not having any rules on posting unit prices. However, in the context of bulk buying inequality, incorporating this measure does not affect the coefficient estimates on income, so it is unlikely that information or education is the driving factor behind the positive relationship between income and package size purchases.

[Table 12 about here.]

7 Conclusion and Future Research

Put conclusion here

Previous research has documented that this "spikiness" has strong implications for measuring consumption inequality (Coibion, Gorodnichenko, and Koustas 2017). Furthermore, this could have implications for inflation measurement as well. Given the prevalence of bulk discounts, if a basket price is computed using the largest package available, this may underestimate the price faced by consumers that cannot afford that package, or buy the same amount in smaller shopping trips. For example, the BLS uses unit-prices when computing the price of the CPI market basket. Therefore, the choice of package size can strongly influence this price. If the prices of different package sizes fluctuate differently, this could systematically affect our measurement of inflation. Further research would be necessary to determine if this is the case and what the magnitude could be.[12]

8 References

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

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review* 99 (4): 1145–77.

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.

Einav, Liran, Ephraim Leibtag, and Aviv Nevo. 2010. "Recording Discrepancies in Nielsen Homescan Data: Are They Present and Do They Matter?" *QME* 8

[12]The BLS specifically deals with multi-unit discounts buy using the "first multiple-unit price" as the price measured. However, the method for dealing with different package sizes is less well defined. Per correspondence with the BLS: "In general, price-per-unit calculations are the standard practice in calculating items with quantity changes. For example if a loaf of bread is $1 and you a second loaf for $0.50, you have two loaves for $1.50, and then depending on the weight of the two loaves, you divide the $1.50 by number of ounces and get a price per ounce. Due to sample rotation, there is no guarantee that any specific item and specific size – toilet paper, olive oil, etc. – will be in the sample at a given time. If it is, and the economic assistant can't find the specified size and has to substitute to a different size, price change is calculated on a per unit basis, as illustrated in the above example." Given that the BLS uses unit-prices to compute market basket prices, the choice of package size could strongly influence the computed price of the market basket.

(2). Springer: 207–39.

Erdem, Tülin, Susumu Imai, and Michael P Keane. 2003. “Brand and Quantity Choice Dynamics Under Price Uncertainty.” *Quantitative Marketing and Economics* 1 (1). Springer: 5–64.

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.

Hendel, Igal, and Aviv Nevo. 2006a. “Measuring the Implications of Sales and Consumer Inventory Behavior.” *Econometrica* 74 (6). Wiley Online Library: 1637–73.

———. 2006b. “Sales and Consumer Inventory.” *The RAND Journal of Economics* 37 (3). Wiley Online Library: 543–61.

———. 2013. “Intertemporal Price Discrimination in Storable Goods Markets.” *American Economic Review* 103 (7): 2722–51.

National Institute of Standards and Technology. 2019. “Uniform Laws and Regulations in the Areas of Legal Metrology and Fuel Quality (Handbook 130).” Edited by Linda Crown, David Sefcik, and Lisa Warfield.

Nevo, Aviv, and Arlene Wong. 2018. “The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession.” *International Economic Review (Forthcoming)*.

Orhun, A. Yeşim, and Michael Palazzolo. 2019. “Frugality Is Hard to Afford.” *Journal of Marketing Research* 56 (1). American Marketing Association.

Pesendorfer, Martin. 2002. “Retail Sales: A Study of Pricing Behavior in Supermarkets.” *The Journal of Business* 75 (1). JSTOR: 33–66.

9 Appendix

9.1 CPI Products and Weights

Table 13 shows the CPI weights of storable product categories that often exhibit bulk discounts. These weights come from the Bureau of Labor Statistics Handbook of Methods, Chapter 17: The Consumer Price Index (Updated 2-14-2018). To be conservative in the importance of accounting for quantity discounts, I capture products that are shelf-stable and storable to varying degrees.

The CPI tries to adjust for quantity discounts in that the "first multiple-unit price is reported for use in the CPI." However, to the extent that there are systematic differences in multi-unit purchases, especially for storable goods, this is likely to overestimate the price index for households that consistently buy in bulk and underestimate the price index for households that buy small quantities.

Full CPI Basket: Includes food and beverages, housing, utilities, apparel, transportation, medical care, recreation, and education.

"Store" Shopping Basket: This basket is constructed to reflect the set of goods commonly found at grocery/supercenter/general merchandise stores. Includes food and alcohol at home (9.024), household furnishings (3.341), apparel (3.343), pet products (0.659), recreational goods (1.5), and other goods (1.634). This excludes categories such as rent, utilities, energy, transportation, health care, education, and services (including food service).

[Table 13 about here.]

9.2 Annual Consumption Analysis (OLS and LASSO)

In order to establish that annual consumption does not substantially depend on income, I first run a basic OLS regression to establish correlations between household covariates and annual purchases. Table 14 provides the OLS results which shows that while income is statistically significant, it is economically miniscule and other household covariates such as household size and presence of children are more predictive of toilet paper consumption than income. Overall, the income coefficients are not statistically different from each other. As a further check, I use a 20-fold cross-validated LASSO regression to select the most predictive covariates and it supports this intuition by choosing household size and other demographics as predictive of consumption and largely shrinks income coeffients to 0.[13]

[Table 14 about here.]

9.3 Comparison with Sales and Coupons

Sales, couponing, and brand choice have gotten considerably more attention in the literature than bulk purchasing. Generally, unit costs are treated as fixed across sizes and consumers choose the quantity to purchase. Sales, coupons, and brands affect the expenditures through the unit price. I have already demonstrated the prevalence of bulk discounts. In order to compare the size of bulk discounting against other savings avenues, I estimate the following regression:

$$log(P)_{krbmt} = \beta_1 log(Q)_{krbmt} + \beta_2 Coupon_{krbmt} + \beta_3 Sale_{krbmt} + \lambda_r + \lambda_b + \lambda_m + \lambda_t + \epsilon_{krbmt}, \tag{8}$$

where P is the unit price paid on shopping trip k at retailer r of brand b in market m at time t. Q is the package size and λ are fixed effects. $Coupon$ and $Sale$ are indicators for whether a coupon was used or the product was on sale.

Table 5 shows that even after controlling for brand and store fixed effects, sales tend to only decrease unit costs by about 5%. In comparison, the same

[13] LASSO results are available upon request.

savings could be realized by purchasing a 33% larger package (e.g. purchasing an 8-pack instead of a 6-pack of toilet paper). Coupons offer slightly larger savings than bulk purchasing, but can still be reasonably outstripped by bulk purchasing. In the case of toilet paper, with common sizes ranging from 4 to 48 units, having a choice over 2, 3, or 4 times larger packages is within most consumer's choice set. In fact, this is likely to be a more available savings method considering that in the Nielsen data, almost 70% of purchases were done without a coupon or sale.

Similar findings hold for tampons, which are shown in Table 6 for completeness.

[Table 15 about here.]

Furthermore, using the Nielsen Retail Scanner Data, when looking at the movement and prices of individual products, sales and price changes are relatively infrequent. Prices are flat over the course of a year with few, short sales, but sales are not a frequent event as they are in the soft drink category (Hendel and Nevo (2013)). Most importantly, even with sales, this often does not change the bulk discounting pattern that larger packages offer lower unit costs.

10 Data Appendix

In order to conduct this analysis, we need a dataset that captures consumer decisions and their available choices. I create the necessary dataset by combining two Nielsen datasets: the Consumer Panel dataset and the Retail Scanner dataset. Nielsen's Consumer Panel dataset contains information on consumer purchases as well as a rich set of demographic variables. Nielsen's Retail Scanner dataset contains information on the weekly volumes of products sold at stores that have agreed to share their data with Nielsen.

10.1 Consumer Panel Data

Households In order to best focus on the consumer choice decision, 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 to purchase toilet paper (e.g. dormitories and barracks). This reduces my sample from about 653,554 households to 644,229 households. I also remove 7 households that have missing residence information.

Products On the products side, there are some products that are obviously miscoded. For example one toilet paper product is reported to have "multi" and "size1_amount" both equal to 36, indicating total package of 1,296 rolls. In reality, only one of those fields was supposed to be 36. I manually correct these discrepancies based on the corresponding product in the Retail Scanner data, because it is likely to have less error. In some cases, I use my best judgement based on similar items offered by the same brand. Overall, 80 products are corrected out of 7,453 total products. I also remove 6 products branded as "to-go" packs since these are unlikely to be used for daily household consumption.

Missing Purchases I then follow the process of Orhun and Palazzolo (2019) to identify missing purchase occasions where a household likely made a purchase, but did not report it. A missing purchase would downward bias my calculation of a household's consumption rate. For example, if a household purchased a 4 pack each month for 3 months, but only recorded the first and last occasion, then their daily consumption would be $8/3 \approx 2.67$ rolls/month instead of the true rate of 4 rolls per month. Given a package size s, I compute the mean and standard deviation of the time until the next purchase. If, the time between purchases given a size s package is longer than the average duration plus 2 standard deviations, then I flag this as a missing purchase occasion.

These "missing" purchases are then used to demarcate when a household is "active", defined as a spell without a missing purchase. Hence, if a household had 1 missing purchase, then they are determined to have 2 active periods. I use these active periods to compute a household's average consumption rate, which is as follows (as per Orhun and Palazzolo (2019)):

$$Consumption_h = \frac{\sum_{a=1}^{A} \sum_{p=1}^{P_a - 1} V_{hpa}}{\sum_{a=1}^{a=A} T_a} \tag{9}$$

where V_{hpa} is the volume of toilet paper for purchase p during active period a. P_a is the total number of purchases made during active period a. All purchases during an active period are included except for the last one because we assume that the purchase made on the final day of an active period is not consumed during that period.

Outliers Even after correcting miscoded products and identifying potential missing purchases, there remain some outliers. In particular, some households purchase excessively large quantities of items, which may indicate they are purchasing for a small business. In other instances, they are inactive, have many missing purchases, or extremely low consumption rates. For my analysis, I remove these outliers, which meet any of the following criteria:

- Excessive Missingness: The household has more than 3 missing purchases and/or has an inter-purchase duration longer than the 99th percentile of all household's maximum inter-purchase duration.
- Inactivity: The household is active for less than 90 days.
- Missing/Insufficient Consumption: Household consumption is below the 1st percentile of the distribution or consumption cannot be calculated.
- Extreme Values: The household purchased a quantity or volume higher than the 99th percentile of the quantity or volume distribution or they spent more than \$50 (inflation-adjusted to 2012) on a single purchase. I also remove any purchases for which a price of \$0 is recorded.

As a result of these conditions, 6.5% of purchases and 28.8% of households are dropped. These exclusion criteria primarily exclude households that make few purchases in this category. Table 16 shows how many observations and households were affected by each condition. Table 17 how the distribution of the sample compares before and after applying these criteria. Overall, these figures mirror those of Orhun and Palazzolo (2019) with the main differences being a result of including the 2015 and 2016 data.

[Table 16 about here.]

[Table 17 about here.]

Finally, for my discrete choice modeling, I restrict the sample to households that purchased a single package. This leaves me with a sample of 3,813,591 purchases across 126,034 households.

10.2 Nielsen Retail Scanner Data

The Nielsen Retail Scanner Data contains weekly pricing, volume, and promotional activity of about 35,000 stores from about 90 retail chains and covers 2006-2016. Each store in the panel reports their weekly quantity sold as well as the weighted average unit price. For a subset of stores, they also report their promotional activity (i.e. if a product was featured or on display).

Because my focus is on the consumer's choice problem, I only extract information related to a consumer's choice set for toilet paper. In order to do this, I combine all movement files of the toilet paper category and then merge those with the Nielsen Consumer Panel data based on the store ID and the week of the shopping trip. This match reduces the possible size of the Consumer Panel because it requires that households shop at the exact stores that have agreed to report their weekly sales to Nielsen.

In examining the consumer choice, I have to make a few assumptions with respect to the Retail Scanner Data.

Assumption 1: *If a product has no sales, it is not in the consumer's choice set. If a product has positive sales, it is in the consumer's choice set.*

In the data, only products with positive sales are recorded, therefore it is difficult to determine if an available product did not sell (a true zero) or if the product was not available (a missing product). Because the data does not distinguish between missing items and those with no sales, I assume that the selection available to consumers is only the set of items with positive sales during the week they shopped at the store.

10.3 Merging Data

I merge the two datasets based on store ID and the week of the consumer's shopping trip. As stated above, I assume that all products with positive sales are in the consumer's choice set and I identify their actual purchase choice by matching the product they chose with the product being offered by UPC. There are some cases where these cannot be precisely matched up and I outline my steps to fix these issues below.

Missing Brand Name Purchases In some cases, the purchase of a brand name cannot be matched by UPC to its corresponding store. In these cases, I simply combine the purchase record from the Consumer Panel with the product assortment offered at that store in the corresponding week. For example, if a consumer records purchasing a 12-pack of Quilted Northern at a particular store in a particular week, but the corresponding store has only recorded sales of other brands (Charmin, Angel Soft, etc.) and maybe other sizes of Quilted Northern (4-, 6-, 8-roll packages), then I add on the consumer's actual purchase to the store assortment. The Kilts Marketing Center has confirmed that this happens occasionally, but they have not been able to ascertain a reason for this occurrence.

Private-Label Purchases In order to maintain anonymity of stores, Nielsen masks the UPCs of private labelled goods in the Retail Scanner Dataset. Since these are commonly purchased, I match private-label purchases between the two datasets based on the following procedure. I record the price, ply, and package size of the item purchased in the Consumer Panel Data. If there is a corresponding item that is the same price, ply, and package size, I record that as a match. This is a conservative matching algorithm and misses some items that might have been purchased at a sale price since the prices would not match up. If there remain unmatched purchases, I follow the same procedure as with the brand-name purchases and append the record from the Consumer Panel to the store's assortment.

Missing Stores Not all individual stores share scanner data with Nielsen or even those that do are not able to be linked to a corresponding store in the Consumer Panel data. In these cases, I impute the choice set available to households by using Assumption 2 below:

Assumption 2: *Stores in the same retail chain and DMA offer the same assortment.*

Given that pricing and inventory policies are often set at a regional or national level, this is likely a weak assumption. Furthermore, the results of DellaVigna and Gentzkow (2017) suggest that pricing is quite similar across stores within the same retail chain. While it is not precisely true that all stores in a chain offer the same assortment of products or offer the exact same price (which is pointed out in DellaVigna and Gentzkow (2017)), I take a conservative approach and generate a "representative" store within a chain by taking the union of the products offered and taking the quantity-weighted average price of those stores for each product. To capture the fact that some chains may coordinate inventories within a region, I generate with "representative" store based on a chain-DMA. For example, if Chain A has multiple stores in Boston and Atlanta, I construct two "representative" stores, one with the set of products offered in the Boston area and the other with the set of products offered in the Atlanta area. If product assortments are set at the national level, there should not be much difference, but this approach is a conservative way to capture the product assortment faced by a consumer.

The results of this matching procedure are recorded in Table 18. I start with the Consumer Panel data that has been cleaned as outlined in the "Consumer

Panel Data" section. We can see that a majority of the sample shrinkage occurs because Nielsen has not assigned certain stores an ID and because some stores that have been assigned an ID do not share their scanner data with Nielsen. Overall, the sample remains large and could be credibly expanded under Assumption 2 below.

[Table 18 about here.]

Collapsing Products For some products, a brand-size does not define a unique product. In these cases, I collapse these items to a composite brand-size with a standardized size, unit cost, price, and ply reflecting the average of the component products. Within a store-week, about 33% of prices are within 1% of the mean price, 50% of prices composing these "composite" products are within 10% of the mean price and 75% are within 25% of the mean price. Overall, price variation is relatively muted within these aggregated brand-size products. These variations are concentrated in certain brands which tend to have more within-brand variation, like Charmin and Angel Soft, as well as the catch-all "Other" category. Allowing for finer gradations of these brands would help reduce this variation. For example, the Charmin "brand" includes their Ultra Soft and Ultra Strong subtypes which currently are combined into a single Charmin product depending on the available size.

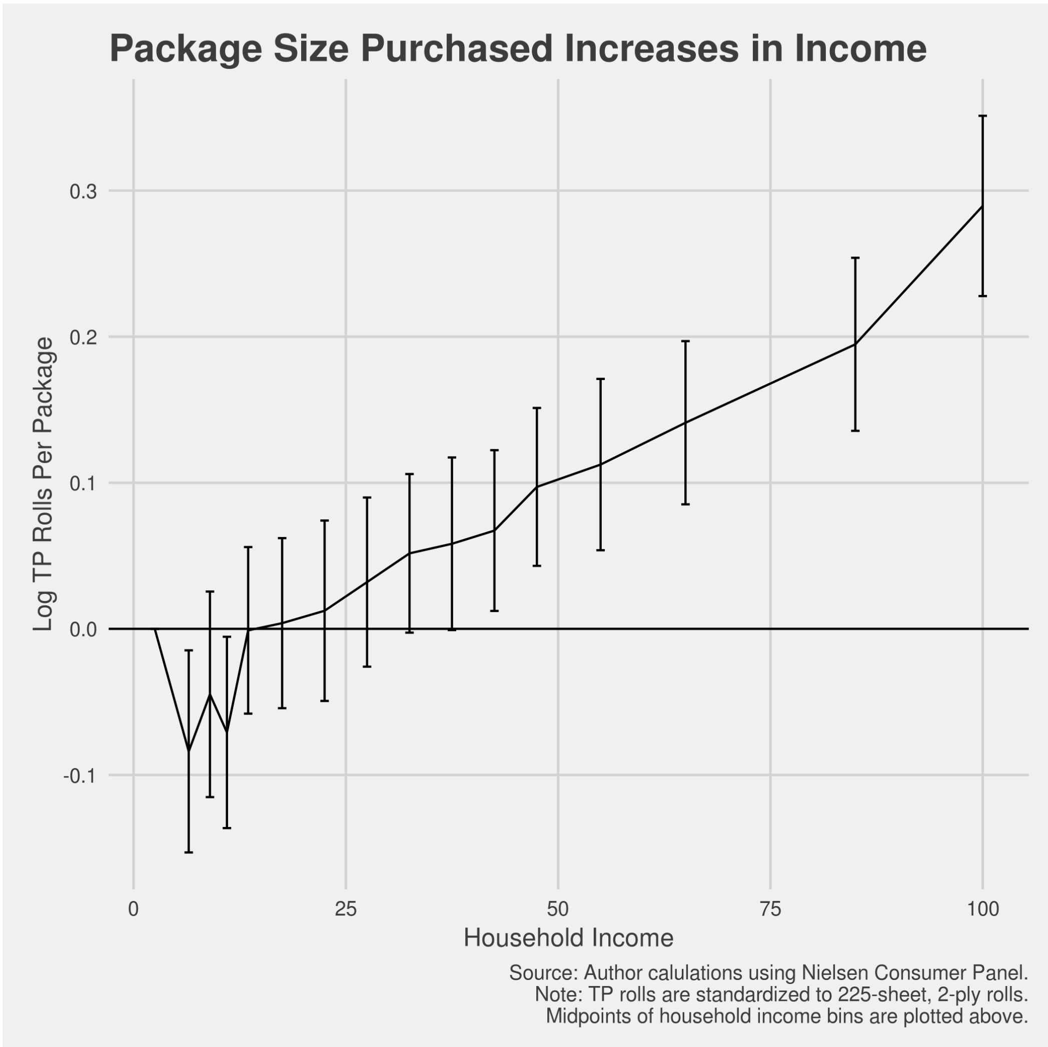
Package Size Purchased Increases in Income
Log TP Rolls Per Package
0.3
0.2
0.1
0.0
-0.1
0
25
50
75
100
Household Income
Source: Author calulations using Nielsen Consumer Panel.
Note: TP rolls are standardized to 225-sheet, 2-ply rolls.
Midpoints of household income bins are plotted above.

Figure 1

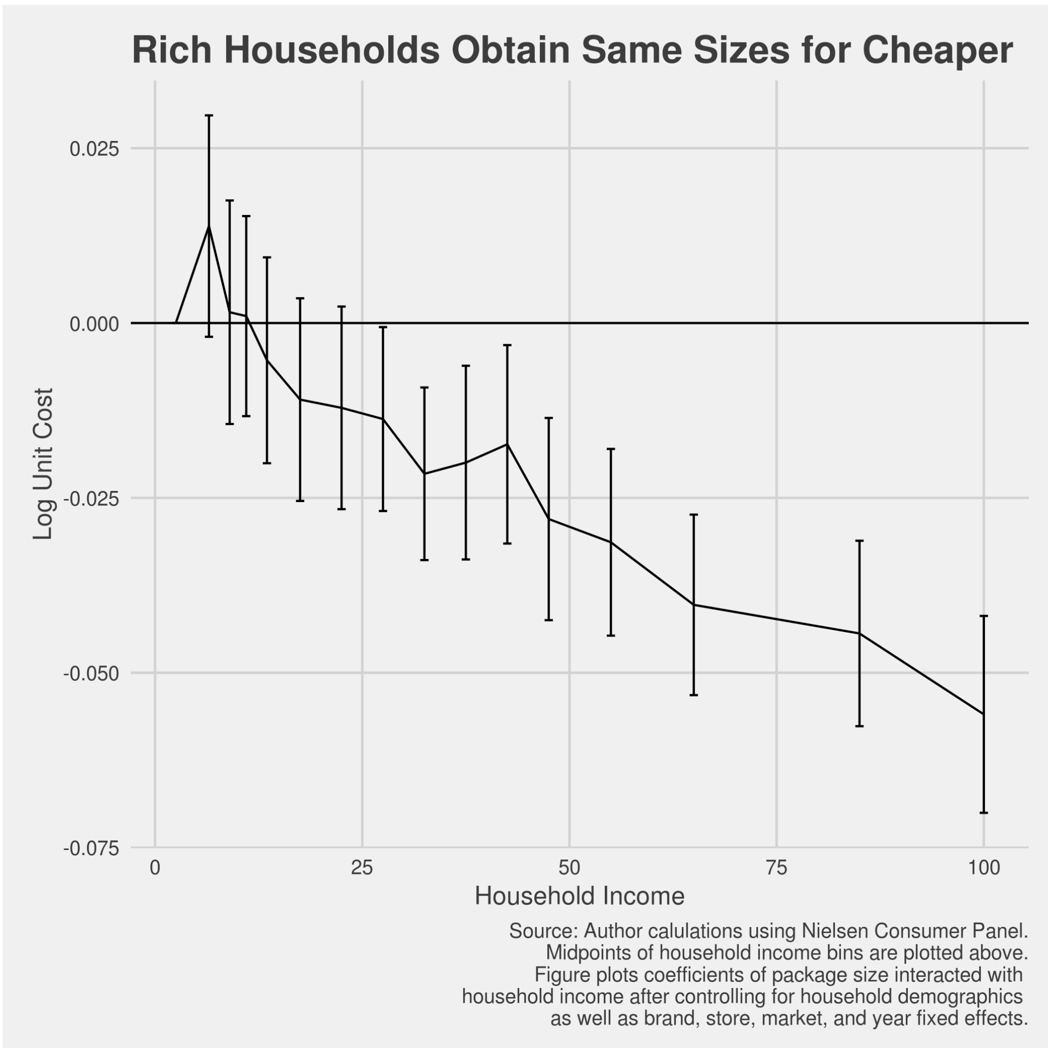
Rich Households Obtain Same Sizes for Cheaper
0.025
0.000
-0.025
-0.050
-0.075
Log Unit Cost
0
25
50
75
100
Household Income
Source: Author calulations using Nielsen Consumer Panel.
Midpoints of household income bins are plotted above.
Figure plots coefficients of package size interacted with
household income after controlling for household demographics
as well as brand, store, market, and year fixed effects.

Figure 2

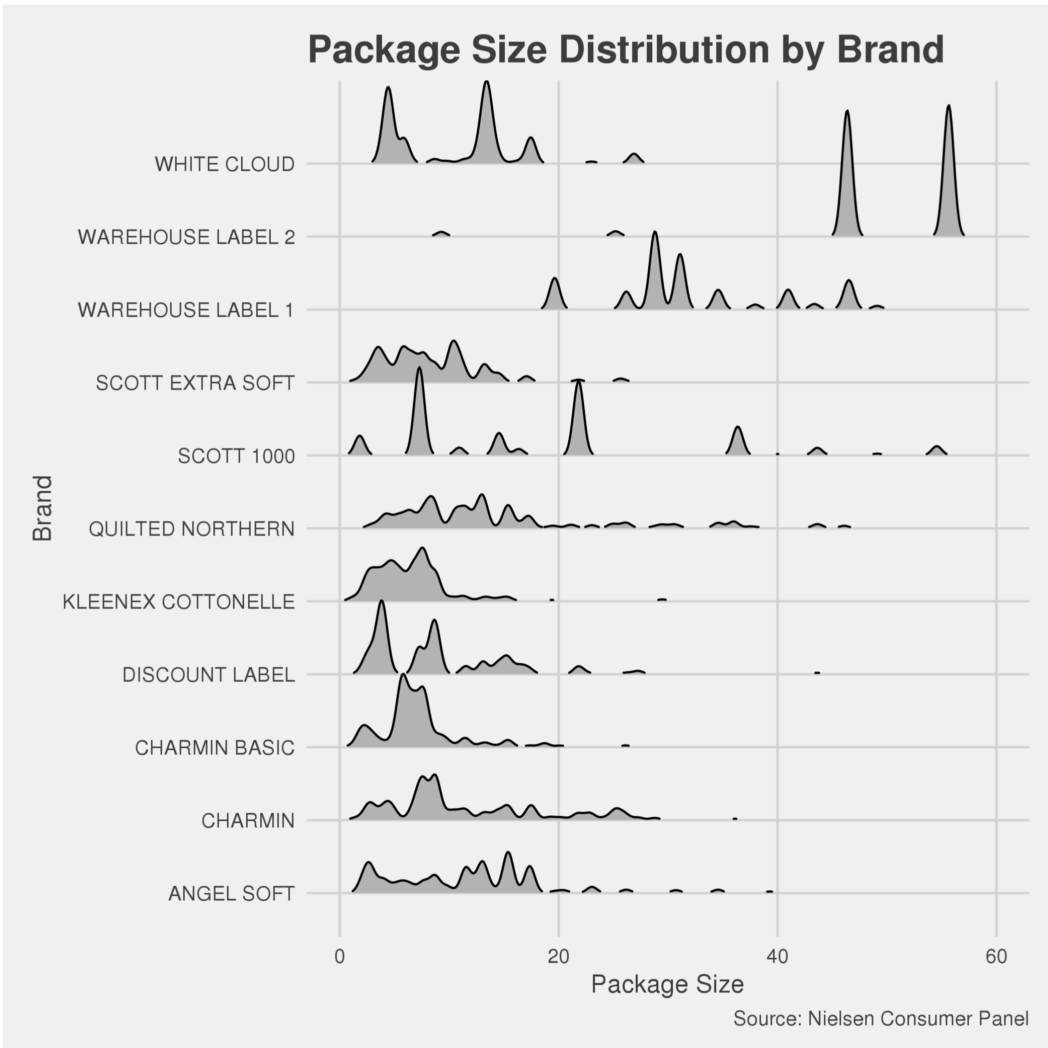
Package Size Distribution by Brand
WHITE CLOUD
WAREHOUSE LABEL 2
WAREHOUSE LABEL 1
SCOTT EXTRA SOFT
SCOTT 1000
QUILTED NORTHERN
KLEENEX COTTONELLE
DISCOUNT LABEL
CHARMIN BASIC
CHARMIN
ANGEL SOFT
Brand
0
20
40
60
Package Size
Source: Nielsen Consumer Panel

Figure 3

Table 1: Prices and Bulk Discounts of Top 5 Brands

Brand	4 Roll Price	6 Roll Discount	12 Roll Discount	24 Roll Discount
Angel Soft	2.11	-0.07	-0.28	-0.33
Charmin	3.60	-0.21	-0.30	-0.38
Cottonelle	3.41	-0.04	-0.20	-0.30
Qltd Ntn	3.48	-0.11	-0.26	-0.35
Scott	3.97	-0.07	-0.21	-0.57

Discounts are per-unit savings.

Table 2: Package Size Statistics by Income Quantile

Income	25th Pctl	50th Pctl	75th Pctl	M	SD	N
<25k	4.160	7.680	13.090	10	8.750	$612,739$
25-50k	4.680	8.290	14.400	11.230	9.820	$1,200,978$
50-100k	5.760	8.730	15.360	12.910	11.330	$1,397,074$
>100k	6.720	11.200	19.640	15.140	12.920	$454,864$

Sizes are in standardized 225-sheet, 2-ply rolls.

Table 3: Toilet Paper Package Size Purchases Increase in Household Income

	Log(Size)			
	(1)	(2)	(3)	(4)
>100k	0.27*** (0.02)	0.29*** (0.02)	0.11*** (0.02)	0.10*** (0.01)
50-100k	0.15*** (0.01)	0.16*** (0.01)	0.07*** (0.01)	0.06*** (0.01)
25-50k	0.07*** (0.01)	0.07*** (0.01)	0.03*** (0.01)	0.02*** (0.01)
Cons. Rate		0.94*** (0.04)	0.73*** (0.04)	0.58*** (0.04)
Time/MSA/Demog. FE	Y	Y	Y	Y
Retailer FE	N	N	Y	Y
Brand FE	N	N	N	Y
Observations	3,665,655	3,665,655	3,665,655	3,665,655
Adjusted R^2	0.08	0.13	0.37	0.45

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are clustered at the market level.
Fixed effects include indicators for year, month, week, MSA, retail chain, and brand. Demographics include household size, housing type, marital status, race, ethnicity, age group, urban/rural indicator, and education.

Table 4: Bulk Discounts Generate Substantial Savings

	Log(Unit Cost)		
	Full Sample		
	(1)	(2)	(3)
Log(Size)	−0.25***	−0.18***	−0.19***
	(0.004)	(0.003)	(0.003)
Year/MSA FE	Y	Y	Y
Brand FE	N	Y	Y
Retailer FE	N	N	Y
Observations	3,372,291	3,372,291	3,372,291
Adjusted R^2	0.19	0.75	0.76

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

Table 5: Bulk Discounts Are Common Across Retailers

	Log(Price)			
Log(Size)	−0.28***	−0.19***	−0.19***	−0.19***
	(0.000)	(0.000)	(0.000)	(0.000)
Brand FE	N	Y	Y	Y
Store FE	N	N	Y	Y
Week FE	N	N	N	Y
Observations	452,223,221	452,223,221	452,223,221	452,223,221
Adjusted R^2	0.17	0.71	0.73	0.77
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 6

	Dependent variable:		
	choice		
	(1)	(2)	(3)
Price (-)	−0.06*** (0.004)	−0.03*** (0.004)	0.02*** (0.004)
Unit Cost (-)	2.18*** (0.04)	2.17*** (0.04)	2.36*** (0.04)
Charmin	1.47*** (0.03)	1.47*** (0.03)	1.31*** (0.03)
Cottonelle	1.82*** (0.04)	1.82*** (0.04)	1.66*** (0.04)
Qltd Ntn	0.25*** (0.03)	0.25*** (0.03)	0.20*** (0.03)
Scott	0.16*** (0.02)	0.16*** (0.02)	0.26*** (0.02)
Other	1.07*** (0.02)	1.07*** (0.02)	1.09*** (0.02)
Small Size	0.81*** (0.04)	0.80*** (0.04)	0.69*** (0.04)
Medium Size	0.51*** (0.02)	0.51*** (0.02)	0.47*** (0.02)
Price/Income Ratio		0.90*** (0.06)	0.71*** (0.06)
Day Supply			−0.01*** (0.0003)
Observations	33,723	33,723	33,723
Log Likelihood	−80,476.24	−80,323.46	−79,741.19
Note:	*p<0.1; **p<0.05; ***p<0.01 Small size is less than 6 rolls. Medium size is 7-12 rolls. Days' supply is the number of standardized rolls in a package divided by a household's average daily consumption.		

Table 7

	Dependent variable:		
	choice		
	(1)	(2)	(3)
p	−4.14***	−3.80***	−3.54***
	(0.50)	(0.34)	(0.21)
unitCost	2.55***	0.96***	0.99***
	(0.04)	(0.02)	(0.02)
brandCharmin	1.37***	1.45***	1.46***
	(0.03)	(0.04)	(0.04)
brandCottonelle	1.41***	1.48***	1.47***
	(0.03)	(0.04)	(0.04)
brandOther	0.44***	0.48***	0.50***
	(0.02)	(0.02)	(0.03)
brandQltd Ntn	0.27***	0.31***	0.33***
	(0.03)	(0.03)	(0.03)
brandScott	−0.94***	−0.96***	−0.96***
	(0.03)	(0.04)	(0.04)
sizeCatmedium	0.76***	0.78***	0.80***
	(0.03)	(0.03)	(0.03)
sizeCatsmall	0.05	0.08*	0.02
	(0.04)	(0.05)	(0.05)
piRatio	0.65***	0.64***	0.64***
	(0.06)	(0.06)	(0.06)
daySupply	−0.02***	−0.02***	−0.02***
	(0.001)	(0.001)	(0.001)
sd.p	0.17	0.15	0.03
	(2.00)	(1.37)	(1.14)
sd.unitCost		0.31***	−0.32***
		(0.03)	(0.03)
sd.daySupply			−0.01***
			(0.001)

Table 8

	Dependent variable:		
	choice		
	(1)	(2)	(3)
p	-2.79^{***}	-2.69^{***}	-2.58^{***}
	(0.07)	(0.07)	(0.08)
unitCost	2.39^{***}	0.90^{***}	1.00^{***}
	(0.05)	(0.02)	(0.02)
brandCharmin	1.34^{***}	1.43^{***}	1.47^{***}
	(0.04)	(0.04)	(0.04)
brandCottonelle	1.69^{***}	1.71^{***}	1.70^{***}
	(0.04)	(0.04)	(0.04)
brandOther	1.04^{***}	1.09^{***}	1.06^{***}
	(0.02)	(0.03)	(0.03)
brandQltd Ntn	0.26^{***}	0.32^{***}	0.30^{***}
	(0.03)	(0.03)	(0.03)
brandScott	0.76^{***}	0.70^{***}	0.60^{***}
	(0.03)	(0.03)	(0.03)
sizeCatmedium	0.61^{***}	0.62^{***}	0.62^{***}
	(0.02)	(0.02)	(0.03)
sizeCatsmall	0.51^{***}	0.54^{***}	0.49^{***}
	(0.04)	(0.04)	(0.04)
piRatio	0.69^{***}	0.68^{***}	0.77^{***}
	(0.05)	(0.06)	(0.06)
daySupply	-0.01^{***}	-0.01^{***}	-0.02^{***}
	(0.0003)	(0.0004)	(0.0005)
sd.p	0.02	0.03	-0.14
	(0.51)	(0.40)	(0.28)
sd.unitCost		0.55^{***}	-0.63^{***}
		(0.03)	(0.02)
sd.daySupply			-0.005^{***}
			(0.001)

Table 9: Credit Access May Help Low-Income Households Buy In Bulk

	Log(Size)			
	(1)	(2)	(3)	(4)
>100k	0.19*** (0.03)	0.20*** (0.03)	0.04 (0.03)	0.04 (0.02)
50-100k	0.11*** (0.02)	0.10*** (0.02)	0.02 (0.02)	0.01 (0.02)
25-50k	0.01 (0.02)	0.02 (0.02)	−0.01 (0.02)	−0.01 (0.01)
Cons. Rate		0.97*** (0.09)	0.70*** (0.07)	0.55*** (0.06)
Credit:>100k	0.07* (0.03)	0.10** (0.04)	0.07** (0.03)	0.07** (0.03)
Credit:50-100k	0.01 (0.02)	0.05 (0.03)	0.05* (0.03)	0.04 (0.03)
Credit:25-50k	0.05* (0.02)	0.06* (0.03)	0.05 (0.03)	0.05* (0.03)
Credit	0.001 (0.02)	−0.04 (0.03)	−0.06** (0.03)	−0.04* (0.02)
Time/MSA/Demog. FE	Y	Y	Y	Y
Retailer FE	N	N	Y	Y
Brand FE	N	N	N	Y
Observations	602,919	602,919	602,919	602,919
Adjusted R^2	0.07	0.13	0.41	0.50

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are clustered at the market level. Fixed effects include indicators for year, month, week, MSA, retail chain, and brand. Demographics include household size, housing type, marital status, race, ethnicity, age group, urban/rural indicator, and education. 'Credit' denotes whether or not the, household used a credit card to pay for a purchase at any point during the year. This is only available for years 2013 and onward.

Table 10: Car Ownership May Help Low-Income Households Buy In Bulk

	Log(Size)			
	(1)	(2)	(3)	(4)
>100k	0.27*** (0.02)	0.28*** (0.02)	0.11*** (0.02)	0.11*** (0.01)
50-100k	0.15*** (0.01)	0.16*** (0.01)	0.07*** (0.01)	0.07*** (0.01)
25-50k	0.06*** (0.01)	0.06*** (0.01)	0.03*** (0.01)	0.03*** (0.01)
Rate		0.93*** (0.05)	0.74*** (0.04)	0.59*** (0.04)
Car Share	0.33*** (0.10)	0.31*** (0.09)	0.23*** (0.09)	0.24*** (0.09)
Time/MSA/Demog. FE	Y	Y	Y	Y
Retailer FE	N	N	Y	Y
Brand FE	N	N	N	Y
Observations	2,595,358	2,595,358	2,595,358	2,595,358
Adjusted R^2	0.07	0.13	0.36	0.44

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are clustered at the market level. Fixed effects include indicators for year, month, week, MSA, retail chain, and brand. Demographics include household size, housing type, marital status, race, ethnicity, age group, urban/rural indicator, and education. 'Car Share' denotes the share of households in the corresponding PUMA that own at least 1 vehicle.

Table 11: Tampon Package Size Purchases Increase in Household Income

	Log(Size)		
	(1)	(2)	(3)
>100k	0.10*** (0.01)	0.03*** (0.01)	0.04*** (0.01)
50-100k	0.05*** (0.01)	0.01 (0.01)	0.02*** (0.01)
25-50k	0.01 (0.01)	−0.01 (0.01)	−0.002 (0.01)
Time/MSA/Demog. FE	Y	Y	Y
Retailer FE	N	Y	Y
Brand FE	N	N	Y
Observations	236,908	236,908	236,908
Adjusted R^2	0.04	0.34	0.43
Note:	*p<0.1; **p<0.05; ***p<0.01 Standard errors are clustered at the market level.		

Table 12: Unit Price Regulations May Help Consumers Make Higher Value Choices

	Log(Size)			
	(1)	(2)	(3)	(4)
>100k	0.27*** (0.02)	0.29*** (0.02)	0.11*** (0.02)	0.10*** (0.01)
50-100k	0.15*** (0.01)	0.16*** (0.01)	0.07*** (0.01)	0.06*** (0.01)
25-50k	0.07*** (0.01)	0.07*** (0.01)	0.03*** (0.01)	0.02*** (0.01)
Law/Guidelines	0.07*** (0.02)	0.06*** (0.02)	0.02** (0.01)	0.01** (0.01)
Cons. Rate		0.94*** (0.04)	0.73*** (0.04)	0.58*** (0.04)
Time/MSA/Demog. FE	Y	Y	Y	Y
Retailer FE	N	N	Y	Y
Brand FE	N	N	N	Y
Observations	3,654,043	3,654,043	3,654,043	3,654,043
Adjusted R^2	0.08	0.13	0.37	0.45
Note:	*p<0.1; **p<0.05; ***p<0.01 Standard errors are clustered at the market level.			

Table 13: CPI Basket Shares

Category	**Basket Share**
Cereals and cereal products	0.370
Processed fruits and vegetables	0.303
Nonalcoholic beverages and beverage materials	0.955
Sugar and artificial sweeteners	0.054
Fats and oils	0.245
Spices, seasonings, condiments, sauces	0.292
Housekeeping supplies	0.847
Personal care products	0.724
Tobacco and smoking products	0.718
Total (all CPI)	4.508
Non-Food Total (all CPI)	2.289
Share of Store Shopping	23.117
Non-Food Share of Store Shopping	11.738

Table 14: Annual TP Purchases

	Average Daily Consumption
5-8k	-0.002 (0.003)
8-10k	-0.01^{**} (0.002)
10-12k	-0.01^{***} (0.002)
12-15k	-0.01^{***} (0.002)
15-20k	-0.01^{***} (0.002)
20-25k	-0.005^{**} (0.002)
25-30k	-0.004^{**} (0.002)
30-35k	-0.01^{***} (0.002)
35-40k	-0.01^{***} (0.002)
40-45k	-0.01^{***} (0.002)
45-50k	-0.005^{**} (0.002)
50-60k	-0.005^{***} (0.002)
60-70k	-0.002 (0.002)
70-100k	-0.004^{**} (0.002)
>100k	-0.003 (0.002)
Multi-Family Home	-0.01^{***} (0.001)
Single-Family Home	0.01^{***} (0.001)
2 people	0.06^{***} (0.001)
3 people	0.11^{***} (0.001)
4 people	0.13^{***} (0.001)
5 people	0.16^{***} (0.001)
6 people	0.18^{***} (0.002)
7 people	0.19^{***} (0.002)
8 people	0.20^{***} (0.004)
9+ people	0.23^{***} (0.01)
Widowed	-0.01^{***} (0.001)
Divorced	-0.01^{***} (0.001)
Single	-0.01^{***} (0.001)
White	-0.002^{***} (0.001)
Child Present	-0.03^{***} (0.001)
Not Hispanic	-0.03^{***} (0.001)
Age	0.001^{***} (0.0000)
Urban	0.0002 (0.001)
College	-0.03^{***} (0.0005)
Observations	481,031
Adjusted R^2	0.13

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
Market and year fixed effects included.
Standard errors are clustered
at the market level. Omitted
categories are the following:
<5k income, mobile homes,
1 person households, married
couples, non-whites

Table 15: Bulk Discounts Provide More Savings Than Sales

	Log(Price)		
	Full Sample		
	(1)	(2)	(3)
Log(Size)	-0.25^{***}	-0.18^{***}	-0.18^{***}
	(0.004)	(0.003)	(0.003)
Sale and/or Coupon	0.02^{***}	-0.23^{***}	-0.23^{***}
	(0.005)	(0.004)	(0.004)
Sale Only	0.04^{***}	-0.05^{***}	-0.07^{***}
	(0.004)	(0.003)	(0.002)
Time/MSA FE	Y	Y	Y
Brand FE	N	Y	Y
Retailer FE	N	N	Y
Observations	3,372,291	3,372,291	3,372,291
Adjusted R^2	0.19	0.78	0.78

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
Standard errors are clustered at the market level.

Table 16: Data Cleaning Steps

Criteria	Obs	Obs %	HH	HH %
Missing 3+ IPD:	168,106	4	6,899	4.6
Max IPD >99th Pct:	21,600	0.5	1,493	1
Cannot calc. consumption:	12,364	0.3	9,031	6.1
Insufficient Consumption:	7,843	0	1,403	0
Active <90 days:	27,475	0.6	14,247	9.5
Abnormal Quantity:	2,088	0	1,189	0.8
Abnormal Volume:	1,913	0	1,307	0.9
Abnormal Price:	6,956	0.2	4,642	3.1
Total HH/Observations:	4,241,992	-	149,272	-
Total Dropped:	520,042	12.3	28,345	19

Table 17: Distribution of Clean and Raw Data

Pctl	Cons. (R)	Cons. (C)	IPD (R)	IPD (C)	Volume (R)	Volume (C)
1st pct	0.04	0.05	0	0	1.82	1.82
25th pct	0.15	0.16	14	14	5.82	5.82
50th pct	0.23	0.23	28	28	9.60	9.09
75th pct	0.34	0.34	59	56	16.58	15.71
99th pct	1.02	0.89	382	305	65.45	55.64
N	149,272	120,927	4,092,720	3,589,017	4,187,109	3,665,655

'R' denotes raw data and 'C' denotes cleaned data.

Table 18: Scanner Cleaning and Merging Steps

Step	Obs	HH
Starting	3,639,516	120,927
Cannot be standardized	3,596,784	120,927
Matched to Scanner	1,153,184	83,102
Single packages	941,702	78,557
Excess Prices	941,701	78,557