Cash-Constrained Households and Product Size*

Tiago Pires and Alberto Salvo

January 17, 2015

Abstract

Storability constrains firms' ability to implement price discrimination, as it enables consumers to separate the timing of purchase from the timing of consumption. Scanner data, however, indicate that significant price differences exist for the same storable good and brand sold in containers of different sizes, with small sizes accounting for a larger share of purchases of low-income households than for their more affluent counterparts. Low-income households therefore tend to be less willing or less able to purchase large storable-good containers typically offered at "bulk prices," to realize savings per unit of consumption, subsequently consuming out of household inventory. We explore one potential explanation for this puzzle: the presence of "cash" constraints, that can be either real or based on heuristics. We describe data patterns that are consistent with cash constraints being present, ruling out alternative explanations for this puzzle. To test for their presence and evaluate implications, we develop and estimate a dynamic discrete choice model that nests the possibility of cash constraints. We find that cash constraints are economically meaningful for low-income households, whereas less relevant for high-income households. Cash constraints further depend on the week of the month. Recognizing that some consumers will only consider offerings at specific price levels is important for manufacturers and retailers when they define their pricing strategies.

1 Introduction

In the summer of 2012, as many European economies were mired in recession, Unilever's regional manager Jan Zijderveld was quoted by Spiegel[1] stating that "Poverty is returning to Europe. As a result, the company has begun offering smaller, less expensive packages so as not to put too great a strain on increasingly limited budgets." The executive further noted that the leading consumer goods firm was transferring such a "smaller package" strategy from its business in the developing

*We are grateful to Javier Donna, Jose Espin-Sanchez, Alwyn Young, Fernando Luco and all participants in the European Winter Meeting of the Econometric Society 2014 for their suggestions. We are also thankful to IRI and particularly Mike Kruger for generously supplying the data. Tiago Pires gratefully acknowledges financial support from Fundação para a Ciência e Tecnologia under the scholarship SFRH/BD/43857/2008. All errors are our responsibility.

[1] "'Poverty Returning': Unilever Cuts Package Sizes in Euro Crisis", Spiegel Online, August 27, 2012 (http://www.spiegel.de/international/business/consumer-goods-giant-unilever-reacts-to-growing-poverty-in-europe-a-852353.html)

countries of Asia, such as Indonesia, where "one-third of Unilever's revenue comes from purchases costing 20 US cents or less".[2]

The smaller packages Zijderveld refers to are often times substantially *more*, not less, expensive per unit of consumption (e.g., a load or serving size), despite the lower price per package. In addition, many or most of Unilever's products are storable for weeks, including home and personal care such as laundry detergent and shampoo, and even (with proper refrigeration) food, such as mayonnaise. The strategy the company seems to be pursuing in both developed and developing markets raises the following question: Why does offering smaller packages and containers at higher prices per consumption unit, usually alongside larger versions of the same brands, help attract or retain households on "limited budgets," presumably those who are most price sensitive? Were it not for some form of constraint, it would seem that households, particularly these price-sensitive ones, would be better off purchasing larger containers of storables over smaller ones, and consuming out of household inventory. Instead, the observation that costly small sizes of brands are popular among European households facing adverse income shocks, or among Indonesia's "emerging middle class," might be construed as evidence that these households care little about prices, a pattern that appears counterintuitive.

In this paper, we explore a mechanism which we argue plausibly explains nonlinear pricing strategies with storable goods and the negative correlation between income and the price per unit of consumption: that of cash constraints, which correlates with household income, cash flow, or wealth.[3] The intuition is as follows. Cash constraints impose an upper bound on the total dollar amount the household is able to spend on the category during the observed purchase occasion. Distance to payday and income, among other socioeconomic observables, may shift this latent dollar threshold. This cash constraint, if it binds, then restricts the set of alternatives that the household considers when choosing which brand-size combination to buy. For example, a low-income household may not consider disbursing \$16.50 on a large 12.5-pound container of Tide, while purchasing the small 3.125-pound container of Tide at a 34% price upcharge per pound. In short, households may choose a small size–with lower total price but higher price per consumption unit–so that they can still benefit from positive consumption when they run out of inventory. The potential gains from stockpiling are therefore lower for cash-constrained households. Ignoring such income-shifting consideration sets—or failing to model such consumer heterogeneity elsewhere—may bias estimates and predictions of economic models.

The presence of cash constraints and forward looking consumers may create purchase acceleration, but cash constraints can still be binding because of (i) persistence of cash constraints, (ii) uncertainty about both the value or timing in which cash constraints bind, and (iii) uncertainty

[2] "Fighting for the next billion shoppers", The Economist, June 30, 2012.

[3] Relatedly, empirical evidence suggests that liquidity constraints partly explain the low adoption rates of beneficial products among the poor in developing countries (Tarozzi et al., 2014).

about the value of purchasing and consuming the good.

Our hypothesis for nonlinear pricing and size proliferation differs from the explanation of most of the existing literature, which ascribes the supply of a brand in different sizes to price discrimination, where a firm's objective in setting high prices per consumption unit for smaller sizes is to meet incentive compatibility (IC) constraints (McManus, 2007; Cohen, 2008). The effects of storability and cash constraints on the IC constraint go in opposite directions. Improved storability (lower storage costs) increases consumers' flexibility and allows consumers to better time their purchases to take advantage of larger containers, making it harder for the IC constraint to be satisfied (Hendel and Nevo, 2013).[4] In contrast, cash constraints soften the IC constraint, because cash-constrained households are less able to purchase large containers.

We start by documenting the aforementioned puzzle using scanner data on all shopping trips and purchases for three product categories—laundry detergent, shampoo, and toothpaste–made by a panel of households over an eleven-year period, beginning January 2001. We provide evidence for nonlinear prices in these product categories and show that low-income households purchase a larger share of small containers. These features of the data explain why the price per pound of products chosen by low-income households is *not* lower than the price per pound for the purchases of higher-income households. We further show that there is negative correlation between the container size and product quality among the purchases of low-income households, but that this correlation is weaker or inexistent among households with higher income. This finding seems consistent with the presence of cash constraints, and less intuitive according to traditional economic models. The sample period we examine, from 2001 to 2011, allows us to evaluate the importance of limited budgets during the 2008/09 financial crisis. We find that during the 2008/09 crisis there was a rise in the market shares of small containers and a fall in the market share of non-small containers. Finally, the data suggests that cash constraints are more likely to bind in the second half of the month than in the first half, particularly for low-income households.

We next develop a model in which a household's decision consists of a two-stage liquidity and brand-size choice process. The first stage determines the set of products (brand-size pairs), within a product category, that the consumer can afford on the observed shopping trip. In the second stage the consumer decides whether to purchase one of those products, taking into account her state, which includes the home inventory level. We introduce storability and cash constraints into a dynamic model of consumer purchasing and inventory, emphasizing that is in the cash constraints that we find it most natural to build in the unobserved heterogeneity. We introduce random shocks both to preferences and cash-constraints, and allow for uncertainty with regard to observable factors

[4] If products were storable at zero cost and consumers were not constrained in any other way, they would arbitrage across different containers available at discriminatory prices within the same brand, so the IC constraint would not be satisfied. Hendel, Lizzeri and Roketskiy (2014) discuss the novel constraints imposed by storability in a model of nonlinear pricing.

that affect preferences and cash constraints. This captures uncertainty about the value and timing in which cash constraints bind and uncertainty on the preferences for purchasing and consuming the good. We then estimate the model using the data for laundry detergent. The estimation of a structural model allows us to test for the prevalence of cash constraints and evaluate their welfare implications. Our work further measures the biases that follow from ignoring such constraints.

The identification of cash constraints is possible by: (1) the availability of longitudinal data, (2) the observed heterogeneity on the likelihood of being cash-constrained, and (3) the variation on the likelihood of being cash constrained within a household due to some observed variable that does not affect the value of purchasing and consuming. The data patterns (and conventional wisdom) suggest that consumers with different income levels differ in the likelihood of being cash constrained, and this likelihood varies with the week of the month, thereby providing the necessary conditions for identification. To illustrate our identification strategy, consider the following example. Suppose all households faced the same storage costs, but rich households were never cash constrained. Then, by taking two households who are identical in all respects but income—namely, a poor household and a rich household—we can identify the separate importance of income and of cash constraints. A "difference in difference" follows from examining the purchasing behavior of both household types during different weeks of the month. In particular, if poor households face a binding cash constraint in the second half of each month, but are not cash constrained in the first half, then the difference in their purchasing behavior in the second half relative to the first reflect the presence of cash constraints (also assume that prices do not vary).

Our results suggest that cash constraints are economically relevant for low-income households, but small for high-income households. Households with income below \$15,000, on average, only consider 84% of the available products and have a probability higher than 15% of considering at most 2 brands in large containers. In contrast, households with income higher than \$55,000 consider nearly 100% of the available products. Cash constraints depend on the time of the month. In particular, they are stronger in the third week when there is a high probability of considering none or a small number of products. Recognizing that certain consumers will consider only offerings with specific price levels is therefore important for manufacturers and retailers when they define markets based on particular price points. We find that ignoring cash constraints underestimates the cross-price elasticity between products and overestimates the cross-price elasticity relative to the outside option. Hence, considering cash constraints is important during merger evaluations and antitrust analysis.

An alternative mechanism to explain the observed negative correlation between income and the price per consumption unit is that households face heterogeneous storage costs, and lower-income households may face higher storage costs. According to our estimates, storage costs are low and therefore we believe cash constraints are a more plausible mechanism to explain the observed patterns. Moreover, it is less plausible that storage costs can explain the within-household choice patterns over week of the month.

We use the estimates of our model to evaluate the consumer welfare implications of cash constraints and understand the trade-off between the amount stored and the preference for "premium" brands over store brands, in the context of storable goods. For instance, cash-constrained households may choose between purchasing the large container of a discount brand and stockpiling, or purchasing a smaller container of a premium brand and not stockpiling. We further calculate how consumer expenditure would differ in the absence of cash constraints. Further, the estimated model allows us to address important questions regarding the supply side: (1) why do premium brands introduce small sizes?, (2) how does the business cycle affect the availability of sizes? (3) how do prices differ across sizes (in light of both IC and cash constraints)?

Related literature. Our work relates closely to a literature on how liquidity constraints may impact individual decisions. Donna and Espin-Sanchez (2014) explore how farmers' financial constraints affect their bids in an auction that allocates water. They develop and estimate a dynamic demand model with liquidity constraints to evaluate welfare changes on moving from an auction system to a system of fixed quotas, finding that the quota raises welfare relative to the auction when liquidity constraints are present. Tarozzi et al. (2014) design a randomized control trial to evaluate the uptake of a health-protecting technology in a poor country under different payment arrangements. The increased adoption when micro-loans were offered, despite a high price for the technology, provides evidence for the presence of liquidity constraints and their importance as a barrier to demand when cash is required upfront. Tarozzi et al. further find that households with lower monthly expenditures were more likely to purchase the technology when micro-loans were available.

The literature on the discriminatory use of package sizes is also related to our paper. McManus (2007) evaluates the distortions arising from differences between marginal benefits and costs when changing product-size menus in the specialty coffee market. He finds that design distortions are large for products not targeted to the highest-demand consumers, but fall toward zero with drink size for products with the largest profit margins. These results therefore support the prediction of "no distortion at the top." Cohen (2008) decomposes the extent to which "quantity discounts" (or bulk pricing) for paper towels can be explained by price discrimination rather than cost differences across sizes. His findings suggest that 35-45% of the observed unit-price variation is consistent with price discrimination. The paper further evaluates counterfactual scenarios in which (1) each brand offers only a small size, and (2) each brand must charge a uniform unit price across all its sizes. Hendel, Lizzeri and Roketskiy (2014) propose a model of nonlinear pricing of storable goods. They show that relaxing the constraints imposed by storability can yield cyclical patterns in pricing and sales. The model provides an explanation for more frequent sales of large packages.

Our paper is related to the literature on the fungibility of money. Hastings and Shapiro (2013) explore how much consumers substitute regular for midgrade gasoline as fuel prices rise compared to when income falls. Using panel microdata on gasoline purchases, they find that the observed response in octane choice to price fluctuations cannot be explained by income effects alone, thus rejecting the null hypothesis that households treat "gas money" as fungible with other income. The authors compare three alternative models of consumer choice that might account for a violation in fungibility: (1) category budgeting, (2) loss aversion, and (3) a salience model. Category budgeting appears to fit the data well, whereas loss aversion and salience both capture some important patterns in the data. In their paper, the focus is on the quality substitution in response to price or income variation, whereas in our paper the focus is on the substitution across different sizes and how these substitution patterns relate to storability and cash constraints

Another strand in the literature studies how macroeconomic conditions impinge on consumers' choices and firms' decisions. Chu and Nevo (2013) document changes in household purchasing patterns during the 2008/09 crisis, with households lowering their transaction prices by increasing their usage of coupons, their frequency of purchasing on sale, their purchase of larger sizes, and their purchase of generic products. These effects are robust to household demographics and were stronger in regions that experienced a larger rise in unemployment. Their findings for the variation in the purchase of large sizes during the 2008/9 crisis contrasts with our findings, but their results also reveal that the share of purchases of products in small containers is larger for low-income households. Despite an increase in shopping intensity, the return to shopping declined during the great recession. Chu and Nevo estimate a high elasticity of substitution between time and goods in home production (consumption declined by one-third less than the fall in market expenditure due to increased home production and time spent shopping), which implies that households were able to partly smoothen consumption through the intra-temporal allocation of time.

Our paper is related to the literature on consideration sets, which considers the possibility that consideration sets form according to a price rule similar to the one we model (e.g., Gilbride and Allenby, 2004). That literature, however, does not explore the interaction between product size and price, and does not evaluate the role of cash constraints. Gilbride and Allenby (2004) investigate the use of screening rules as part of a discrete-choice model. Their framework assumes that alternatives that pass the screen are evaluated in a manner consistent with random utility theory, whereas alternatives that do not pass the screen have a zero choice probability. Illustrating the model in the context of consumer preferences for a new camera format, they find that screening rules restrict the set of alternatives that are evaluated for final selection.

Our use of week of the month to identify cash-constraints relies on some of the ideias and empirical evidence of literature evaluating the effects of paycheque day/week on the timing of purchases and consumption. Melvin (2006) finds that household consumption is excessively sensitive to paycheque arrival using data from the UK's Family Expenditure Survey. The excess sensitivity of consumption to paycheque arrival cannot be explained by common shocks to all households, but can be accounted for the presence of liquidity constraints as measured by wealth and age. Huffman and

Barenstein (2005), using the same data as Melvin, also find evidence that consumption spending declines between paydays, and jumps back to its initial level on the next payday. Shapiro (2005) finds a preference for immediate consumption with average caloric intake declining by 10 to 15 percent over the food stamp month. These results are not reconciled with exponential discounting but can be explained by hyperbolic discounting.

Finally, our work connects to the literature on switching costs and brand loyalty. For Unilever's manager Jan Zijderveld, hanging on to some consumers in some form, in the hope that they trade back up to more profitable within-brand offerings (e.g., larger sizes or higher prices) when they again enjoy positive income shocks, may be strategic. Losing these consumers to store brands and generics today (which are given generous shelf space in retailers such as Aldi, Lidl, etc) means it may be more costly to win them back in the future. In the presence of brand loyalty and cash-constrained households, firms may therefore face incentives to cut prices or sizes during recessions, as they do not want to lose loyal consumers only to spend resources attempting to recover them during booms.

2 Data Description

We use household panel data and aggregate store-level data collected by Information Resources Inc. (IRI) during eleven years, beginning January 1, 2001.

The household panel data are drawn from two Behavior Scan markets in the US: Eau Claire, Wisconsin, and Pittsfield, Massachusetts. This dataset contains information for all shopping trips and the complete purchase history for several product categories, including the total expenditure on each shopping trip, the universal product code (UPC) for each of the purchased products, the number of purchased containers, and the total amount of dollars spent on each product.

The store-level data cover 50 IRI markets and include the average price charged, the aggregate quantity sold, and the promotional activities for each product at each store during each week. The data record two types of promotional activities: feature and display. The feature variable measures whether the product was advertised by the retailer, and the display variable captures whether the product was displayed differently than usual within the store that week.

The appendix describes the combined dataset and details how we cleaned the raw data for each product category we examine, namely liquid laundry detergent (laundry detergent hereafter), shampoo, and toothpaste. An observation in the final sample is a shopping trip. The final sample for laundry detergent consists of 537,776 shopping trips by 2192 households over the eleven-year period (560 weeks).[5]The data records purchases across 36 stores in total. Across store-week pairs in the sample, the median number of brands and UPCs, respectively, purchased are 5 and 7 in laundry

[5]For shampoo consists of 240,166 shopping trips by 1,256 households and for toothpaste consists of 448,521 shopping trips by 1,626 households.

detergent, 3 and 3 in shampoo, and 4 and 6 in toothpaste. For the estimation of the structural model we consider a subsample with 39,277 shopping trips. We convert all prices, including household income and expenditure, to January 2006 US dollars using the national CPI for all goods (source: BLS).

Figure 1 plots the income distribution for households who purchase in each of the three product categories.[6] The bell-shaped distributions include a substantial mass of households earning low incomes. The presence of low-income shoppers in the sample is important, as these households are more likely of becoming cash-constrained at some point in time, such as during the month. If cash constraints depend on income, the observed dispersion in the income distribution will provide variation in the extent to which cash constraints bind across households.

Figure 2 describes the within-household quarterly variation in total shopping expenditure and category shopping expenditure for households who purchase in each of the three categories. The panels are indicative of large and seasonal variation in expenditure. Expenditures do not exhibit a specific trend. A decline is observed during the 2008 crisis for toothpaste, but, interestingly, the expenditures in laundry detergent slightly rose during the crisis, which can be explained by some of the facts described later.

Figure 3 summarizes the within-brand mean price per pound across different size ranges in each product category.[7] The nonlinearity of pricing schedules is evident, with the smaller containers being sold at higher prices per consumption unit. In particular, the price per pound for small-size products is substantially larger than the price per pound for medium- and large-size products. The evidence of nonlinear pricing is important, because this pricing strategy may not be possible to implement for storable goods, unless there are some frictions that make consumers self-select into different groups.

Table 1 display some summary statistics, with an emphasis on laundry detergent. The mean price per pound for the products available during a shopping trip is \$1.06/lb. Most households go shopping once or twice a week. The mean time elapsed between trips is 5 days, but the mean time elapsed since the last purchase of liquid laundry detergent is 8 weeks. Total average expenditure per trip is \$61, and the proportion of trips to stores visited by the household at least once in the preceding 12 weeks is 96%, suggesting that households know the stores they visit fairly well.

[6] The demographic data was updated in 2005, 2007, and 2011, and was current as of the date of the update. The data indicate that income is quite persistent, suggesting that the absence of updates in the remaining years (which we impute as explained in the appendix) does not lead to significant measurement error. This data limitation, however, does not allow us to observe the full time path of income.

[7] In the laundry detergent category, products with size lower than 4lb were assigned to the small size, products with size between 4lb and 8lb were assigned to the medium size, and products with size greater than 8lb were assigned to the large size. In the shampoo category, products with size lower than 0.8lb were assigned to the small size, products with size between 0.8lb and 1.5lb were assigned to the medium size, and products with size greater than 1.5lb were assigned to the large size. In the toothpaste category, products with size lower than 0.3lb were assigned to the small size, products with size between 0.3lb and 0.45lb were assigned to the medium size, and products with size greater than 0.45 were assigned to the large size.

A household purchased, on average, nearly 32 units of liquid laundry detergent and 4 different brands. The median household has total income averaging $40,000 over time, and is composed by 2 members. Tables 3, 13, and 16 explore the importance of household heterogeneity on choices with regard to different household characteristics. These tables reveal important heterogeneity with income, household size, and age of the head of the household.

Table 2 reports market shares (by volume and by revenue), mean prices and the frequency of promotional activities in the laundry detergent category for the 24 leading brand-sizes (by volume share) in the sample. Volume shares range from 12.64% to 0.05%. We again see that prices per pound exhibit substantial variation both across brands and across sizes within brand. There is heterogeneity in the frequency of in-store displays (ranging from 2% to 12%) and feature ads (between 3% and 13%) across the reported brand-sizes. The top 10 products account for roughly two thirds of volume purchased and the top four sellers (Procter&Gamble, Church&Dwight, Dial and Lever Brothers) enjoy a combined share in excess of 80%. Naturally, market concentration at the household level is even more pronounced since households differ in their preferred brands.

3 Preliminary Analysis

This section documents household purchasing patterns and evaluates whether these patterns suggest the presence of limited budgets and cash-constrained households. The evidence presented here will be used to construct the model described in the following sections. Our arguments are based on the idea that goods are storable and the same brand is sold in different containers' sizes.[8] Consumers could therefore make savings over time through stockpiling by buying the products with the lowest price per unit of consumption. Moreover, the validity of our arguments relies on low storage costs and no perishability. Hence, in our analysis we choose 3 product categories–laundry detergent, shampoo, and toothpaste–that satisfy these conditions. We restrict our attention to non-food items, because perishability and large storage costs associated with refrigeration can be a concern with food items.

Figure 5 plots the share of purchases of each container-size by household's income. This figure shows that the share of purchases of small containers is larger for low-income households and the share of purchases of large containers is greater for high-income households in the laundry detergent and toothpaste product categories. As these goods are storable and there is a negative correlation between container size and price per unit of consumption, this result may seem counter intuitive. Yet if consumers have limited budgets, the described pattern can be explained by the lower total price of small containers, which makes them affordable for low-income households, whereas large

[8]Unfortunately for our analysis in the shampoo category some top brands (i.e. brands with high market share) do not sell all the sizes. This makes our analysis more difficult for this product category and may help explain why some of the patterns in this product category slightly differ from the patterns in the other two product categories.

packages may be unaffordable. In particular, the combination of cash-constraints and random-shocks to consumption or income may explain the purchase of products with higher price per pound but lower per container spending, because the purchase of these options may be the only way of ensuring a positive consumption when a cash-constrained household is hit by a shock that makes the household run out of inventory.

A positive correlation between income and the purchased size of laundry detergent is further suggested by the first column in table 4, which display OLS estimates of the effects of income on size choice. These columns reveal that households with higher income are less likely of buying larger sizes of laundry detergent, even controlling for household and shopping trip characteristics. For shampoo and toothpaste, our estimates suggest a nonsignificant relation between income and the size purchased, again suggesting that, against conventional wisdom, low-income households do not make savings through stockpiling by buying the products with the lowest price per unit of consumption. As the total price of laundry detergent, particularly for the large pack-sizes, is usually larger than the total price for the other two product categories, the stronger effects found for laundry detergent can also be interpreted as evidence of the presence of cash constraints, because those are more relevant for more expensive product categories.

If cash-constraints are the reason for low-income households not buying large containers, then the time of the month can be important in the choice of the size purchased. For instance, if shocks to consumption are random and utility of consumption is such that households prefer to avoid zero consumption, then an unanticipated negative shock to consumption could make a household run out of inventory and therefore yield the purchase of small containers, because this is the only affordable option when cash-constrained, even though it is also the more expensive option in terms of price per unit of consumption.[9] The described situation is more likely to happen later in the month, because payday tends to be in the beginning of the month and therefore cash constraints are likely to bind then. Column 3 in table 4 suggests that this happens at least for laundry detergent. We find that households with income below the median income are less likely of buying a larger container of laundry detergent and this likelihood of buying a larger container falls over the month. For the other two product categories we do not find an effect of period of the month on the size purchased, which may be explained by less frequent purchases of these product categories. Less frequent purchases may allow households to timing their purchases for periods in which they are less cash-constrained. That is, for less frequently purchased goods consumers may be willing to accept a zero consumption for a short-period of time and thus the unanticipated shocks become less relevant.

The information for income in our data was updated only in 2005, 2007 and 2011. So, it is likely that there is a measurement error in the measure that we are using, which can bias our results.

[9] A similar argument can be used regarding unantecipated shocks to the available cash.

Moreover, this reduces the variation that we can observe in income over time, making almost impossible to use fixed effects to control for household characteristics as these will be collinear with income. In order to check the robustness of our results to these flaws, we evaluate the effects of income on the purchased size, using the household's quarter expenditures (excluding the product category under analysis) as a proxy for income and controlling for household fixed effects.[10] The fourth column in tables 4 and 5 for each product category reports the results for one of these regressions. The results seem to be similar, even stronger, to the results obtained using our measure of household income, suggesting some robustness of our results.

Figures 6 to 8 display the share of purchases of each container-size by income, controlling for family size. These graphs show that our results are robust to family size and thus the larger share of purchases of small containers by low-income households is not driven by a correlation between income and family size

Fact 1: *The share of purchases of small containers is larger for low-income households.*

The period of our sample includes a major recession, thus providing significant variation in income and wealth that we can explore to evaluate the effects of income on consumers' choices. Figure 9 shows the quarterly evolution of real GDP and container-size market shares. These figures show that until 2007 the share of large containers was rising over time for the 3 product categories (particularly for laundry detergent and toothpaste), whereas the share of small containers was either falling (detergent and toothpaste) or stable (shampoo). In late 2007 those trends are interrupted with a rise in the market share of small containers and a fall in the market shares of large and medium containers. The rise in the share of small containers–particularly sharp for laundry detergent and toothpaste–and the fall of either medium or large containers became even stronger during 2008. Those trends stop only in 2009. These patterns seem to hold across the 3 product categories, even though there were differences in the form of adjustment. For instance, for toothpaste the rise in the share of small containers happened essentially through a fall in the share of medium containers, whereas for laundry detergent the fall in the share of large containers is sharper. This reveals that during the 2008/2009 crisis the share of small containers rose and the share of non-small containers fell, which may suggest that the purchased containers' size varies over the business cycle. For laundry detergent, our data further reveals a large drop in the total quantity purchased after May 2008. This is due to a large fall in the size of the products purchased. Our data also suggests that some large containers disappeared from the market (or consumers were not buying them). Overall, purchases of small (large) containers look counter (pro) cyclical and thus the transaction price per

[10] We use household grocery expenditures to proxy for time-varying shocks to household income because existing literature shows that food expenditure responds to variation in income in the cross-section and over time, predicting about 40% of the cross-sectional variation in total expenditure (Skinner 1987) and responding significantly to shocks to current and future household income (Stephens 2001, 2004; Jappelli and Pistafferi 2010). See Hastings and Shapiro (2013) for an application and discussion of this proxy.

pound that each household pays can rise during a recession on account of this mechanism alone.

Fact 2: *During the 2008/2009 economic crisis there was a large rise in the market share of small containers and a fall in the market share of non-small containers.*

Figure 10 plots the average transaction price per pound with respect to household income. This figure does not suggest either substantial lower transaction prices per pound for low income households or a positive correlation between income and the average price per pound. Households with income lower than $20,000 indeed pay one of the highest average price per pound for laundry detergent. These results are surprising, because conventional wisdom and results in the empirical and theoretical literature often suggest that high-income households are less price sensitive and thus we would expect that the mean transaction price would rise with income (Nevo 2001). Moreover, these findings cannot be explained by small variation in the prices of the different available alternatives because there are large price differences across products as illustrated in figure 4.

Figure 11 displays the average transaction price with respect to the number of TV's that households hold. If the number of TV's is a good proxy for income, then this figure strengths the previous results.

Facts 1 and 2 can help to explain why we do not find a lower transaction price per pound for low-income households: low-income households are more likely to purchase small containers, which are more expensive, and this increases their transaction prices per pound. In order to study this potential explanation with more detail, figure 12 plots the average transaction price per pound stratified by income and container size. This figure shows that some of the aforementioned effects disappear or are weaker when controlling for the size purchased. This suggests that the high transaction price per pound for low-income households is partly due to the purchase of a larger share of small and medium containers, which are more expensive.

Fact 3: *The transaction price per pound for low-income households is not lower than the transaction price for households with higher income.*

Figure 13 reports the ratio of the share of purchases of premium brands during purchases of large containers to the share of purchases of premium brands during purchases of non-large containers by household's income. The figure shows that the share of purchases of premium brands is larger during purchases of non-large containers for households with income below $20,000. The displayed ratio, however, rises or even becomes greater than 1 for households with larger income. These results suggest that there is a negative correlation between quality and size for low-income households but this correlation is weaker or does not hold for households with higher income.

Fact 4: *The choice of quality and container's size is negatively correlated for low-income households, but this negative correlation is weaker or does not hold for households with higher income.*

Figures 14 to 19 show that purchase and shopping behavior tend to vary cyclically over the month. In particular, purchases, expenditures and transaction prices tend to be higher in the first half of each month. These patterns are expected in the presence of cash constraints, because cash constraints are more likely to bind later in the month with a longer distance from payday.

Fact 5: *Purchase and shopping behavior varies cyclically over the month.*

One potential explanation for the results described in this section is that households with low income are time inconsistent or have hyperbolic discount. Yet, we do not believe this is the explanation for the results that we found, because it would require that consumers do not learn and repeatedly underestimate their consumption. As we observe households for eleven years, we believe that this is unlikely for a so large period of time.

Another potential mechanism that helps explain the negative correlation between income and share of purchases of small containers is that households face heterogeneous storage costs and lower income households face higher storage costs. Yet the large magnitude of the effects that we found can only be explained by this mechanism if storage costs are large. We, however, do not believe that storage costs are extremely large in the categories that we analyzed. Moreover, in our structural model, we will allow for the possibility of storage costs and therefore we can test whether storage costs are high for laundry detergent.

The presence of cash constraints and forward looking consumers may create purchase acceleration, but cash constraints can still be binding because of (i) persistence of cash constraints, (ii) uncertainty about both the value or timing in which cash constraints bind, and (iii) uncertainty about the value of purchasing and consuming the good.

4 Model

This section presents a dynamic model of consumer choice for a storable good. To account for some of the patterns described in Section 3, the model nests the possibility that the consumer faces a constraint, at the time of purchase, which restricts the set of products that she is willing or able to consider purchasing within the category. We model the constraint as a maximum out-of-pocket amount that the consumer is willing to pay to purchase one container on the shopping occasion, regardless of the container's size in consumption units, and thus label this a "cash" constraint. We subsequently discuss possible interpretations of this container-price based cutoff, whether it is based on real liquidity constraints or on heuristics. The introduction of a "cash" constraint in the model is motivated by the "unexpected" data pattern of low-income households purchasing a larger share of small containers and the consequent inexistence of differences in the prices per pound paid by households with different income. Our model nests other possibilities for the purchase of a small container by low-income households, and therefore one of the advantages of our model is to test the hypothesis that consumers are cash constrained against alternative hypothesis. Our model, however, cannot nest all possibilities for the purchase of small containers by low-income households. We will discuss for some of these alternatives why we believe that they are not the explanation for the patterns observed in the data.

Choice is modeled as a two-stage process where consumer i at time t first establishes the set of products that she can afford, Ω^*_{it}, and then decides whether to purchase one of the products in this restricted choice set. Thus, not all products that are available at the time of purchase, Ω_t, are *a priori* substitutable for the consumer. Products indexed by j are characterized by brand-size combinations and include the no-purchase option $j = 0$. Let p_{jt} and q_j denote, respectively, the price per container and units of consumption per container (e.g., lb/container) for alternative j; then, the price per consumption unit is p_{jt}/q_j ($/lb in the example). We model the restricted choice set as

$$\Omega^*_{it} = \{ j |\, p_{jt} \leq p^*_{it} = \theta x_{it} + \varsigma_{it} \}$$

where the out-of-pocket cutoff p^*_{it} shifts in ways both observed, x_{it}, and unobserved, ς_{it}, to the analyst, but observed to the consumer at time t. Variables x_{it} may include week-of-month dummies (e.g., capturing distance from payday at month-end), month-of-year dummies, income, and proxies for "disposable" income. We assume that the consumer has perfect foresight over some x_{it} variables (e.g., week of month evolves deterministically and cycles regularly) but not others (e.g., the macroeconomy).

Hence, a consumer facing a cash shortfall this period—i.e., facing low x_{it} or low ς_{it}—may choose between purchasing a small container (and preserving liquidity), or not making a purchase this period in anticipation of purchasing a large container next period, when her out-of-pocket constraint is less severe. Relative to a large container j', a small container j is typically characterized by a low out-of-pocket expense but a high price per unit of consumption, i.e., $p_{jt} < p_{j't}$ and $p_{jt}/q_j > p_{j't}/q_{j'}$. The constrained consumer would be trading off a high-price purchase to meet consumption today against running out of inventory today but purchasing at a lower price per consumption unit tomorrow. Similarly, a consumer with a large positive draw ς_{it} may purchase to inventory today to take advantage of a low price per unit of consumption, in the event she is constrained tomorrow.[11]

Consumer i chooses product $j \in \Omega^*_{it} \subset \Omega_t$, where the flow value from purchasing j at time t is

[11] This model also allow us to discuss the fungibility of money and the trade-off of stocking the good versus stocking the cash. Our constraint may indeed reflect a preference for stocking cash instead of stocking the good. It is therefore important to understand why consumers cannot simply stock the good or why inventory costs of detergent may exceed inventory costs of cash. Moreover, our estimated model should be able to tell us how often each consumer is constrained, to the extent she shifts in and out of small containers.

given by

$$
\begin{aligned}
u_{ijt} &= \alpha_i p_{jt} + \gamma a_{jt} + \xi_{ij} + \epsilon_{ijt} \\
&= \delta_{ijt} + \epsilon_{ijt}
\end{aligned}
$$

where a_{jt} are observed non-price attributes, ξ_{ij} is an unobserved idiosyncratic taste for product j that varies across consumers and can be a function of product characteristics, and ϵ_{ijt} is a random shock to consumer choice.[12] The value associated with a no-purchase is

$$u_{i0t} = \epsilon_{i0t}$$

We assume $\{\epsilon_t\}_{t=0}^{\infty}$ are independent and identically distributed extreme value type I. Let indicator variable d_{ijt} denote the purchased product, thus equal to 1 if consumer i chooses alternative j at time t and 0 otherwise. In each purchase occasion a consumer can buy at most one "inside" product, $\sum_{j \geq 0} d_{ijt} = 1$. In our framework, we do not attempt to model the choice of store we observe in the data. The timing and incidence of shopping trips are exogenous.[13]

The good is storable and thus the quantity not consumed in a given period remains stored as inventory. Consumer i obtains per-period utility of consumption $U_i(C_{it})$, where C_{it} is period t's total quantity consumed, in units of consumption, across all varieties (brand-sizes j) for the category in question. As in Hendel and Nevo (2006), we model products as perfect substitutes in consumption, i.e., utility differs across products at the moment of purchase but not during consumption.

Finally, as the utility from consumption does not depend on which brands are in storage, we can define inventory I_{it} as the total quantity stored, in units of consumption, across all varieties in the category. The evolution of inventory is described by:

$$I_{it} = I_{it-1} + \sum_j d_{ijt} q_j - C_{it}$$

where I_{it-1} is inventory at the start of the current period. There is disutility from holding inventory, with consumer i incurring cost $T_i(I_{it})$ to store composite quantity I_{it} for goods in the category.[14]

In sum, each period consumers choose the brand to buy, how much to buy, and how much to

[12]Price and non-price attributes are store specific. The idiosyncratic tastes and the random shocks can also be store specific. Store subscript is omitted only to simplify notation.

[13]This assumption implies that the intention to purchase detergent does not cause consumers to go shopping. This assumption is supported by the evidence that consumer decision making usually occurs in store (Hoch and Deighton, 1989, Dreze, Hoch and Purk, 1994). According to Seiler (2013), this assumption is also reasonable because detergents are a small fraction of the total expenditure on the typical shopping trip and there are small effects, if any, of store-traffic promotions on individual items.

[14]In the estimation we assume that storage costs are equal for all households with the same inventory holdings. We can relax this assumption, but discussion of the identification of the parameters of the model is easier with this assumption.

consume. We assume that the purchased quantity is a choice of size, $q_{it} = \sum_j d_{ijt} q_j$ (i.e., consumers choose the container size, not how many containers).

Consumers are forward looking and maximize the present expected value of future utility flows. Let s_t be the state at time t and ρ be the discount factor. To simplify notation, we omit subscript i hereafter. The consumer's problem is:

$$\begin{aligned}
V\left(s_1\right) &= \max_{d_{jt}\in\{0,1\}} \sum_{t=1}^{\infty} \rho^{t-1} E\left[U\left(C_t\right) - T\left(I_t\right) + \sum_j 1\left[p_{jt} \leq \theta x_t + \varsigma_t\right] d_{jt}\left(\delta_{jt} + \epsilon_{jt}\right) \middle| s_1 \right] \\
s.t. &: I_t \geq 0 \\
&: C_t \geq 0 \\
&: I_t = I_{t-1} + \sum_j d_{jt} q_j - C_t \\
&: \sum_j 1\left[p_{jt} \leq \theta x_t + \varsigma_t\right] d_{jt} = 1
\end{aligned}$$

The state variables include inventory, week of the month, income and other household characteristics (e.g., family size), consumption rate, prices, product displays, and feature ads. Week of the month evolves deterministically and our model implies that inventory is also deterministic. We further assume that income, household demographics and consumption rates are time invariant.

Consumers face uncertainty about random future prices, displays, and feature ads. We make the following assumptions

Assumption: *Prices, product display, and feature ads are independent and identically distributed over time.*

We also make the following assumption regarding the random shocks ς

Assumption: *The random shocks ς are independent of past actions and of past and current realizations of the other state variables with known distribution F_ς independent and identically distributed over time and across households.*

The function $V\left(s_t\right)$ is the value function and is the unique solution to the Bellman equation characterized by

$$V\left(s_t\right) = \max_{d_{jt}} \sum_j d_{jt} 1\left[p_{jt} \leq \theta x_t + \varsigma_t\right] \left\{\left(\delta_{jt} + \epsilon_{jt}\right) + M\left(s_t, j\right)\right\}$$

where

$$M\left(s_t, j\right) = U\left(C_t\right) - T\left(I_t\right) + \rho E\left[V\left(s_{t+1} | s_t, d_t\right)\right]$$

Our assumptions imply that current actions affect utility only through inventory. Hence, the expectation of V given state and current behavior–denoted by V^e–is a function of current price and inventory. That is,

$$V^e\left(I_t, p_t, a_t\right) = E\left[V\left(s_{t+1}\right)|I_t, p_t, a_t, d_t\right]$$

$$\begin{aligned}
&= \int \max_{d_{jt+1}} \sum_j d_{jt+1} 1\left[p_{jt+1} \leq \theta x_{t+1} + \varsigma_{t+1}\right]\left\{\left(\delta_{jt+1} + \epsilon_{jt+1}\right) + M\left(s_{t+1}, j\right)\right\} dF\left(\epsilon\right) dF\left(\varsigma\right) dF\left(a\right) dF\left(p_{t+1}|p_t\right) \\
&= \int \ln \sum_j \exp(1\left[p_{jqt+1} \leq \theta x_{t+1} + \varsigma_{t+1}\right] \times \\
&\quad \times \left\{\delta_{jt+1} + U\left(C_{t+1}\right) - T\left(I_{t+1}\right) + \rho V^e\left(I_{t+1}, p_{t+1}\right)\right\}) dF\left(\varsigma\right) dF\left(a\right) dF\left(p_{t+1}|p_t\right)
\end{aligned}$$

Let $\{p_{kt}\}_{k=1}^J$ be an ordered sequence of prices observed in period t, such that

$$p_{1t} \leq p_{2t} \leq ... \leq p_{Jt}$$

The unconditional probability that consumer i chooses product $j > 0$ at time t is

$$\begin{aligned}
\Pr\left(d_{jt} = 1\right) &= \sum_{k=1}^{J-1} \Pr\left(d_{jt} = 1|p_k \leq p_t^* < p_{k+1}\right) \cdot \Pr\left(p_k \leq p_t^* < p_{k+1}\right) + \\
&\quad + \Pr\left(d_{jt} = 1|p_t^* \geq p_J\right) \cdot \Pr\left(p_t^* \geq p_J\right)
\end{aligned}$$

noting that $\Pr\left(p_k \leq p_t^* < p_{k+1}\right) = F\left(p_{k+1} - \theta x_t\right) - F\left(p_k - \theta x_t\right)$ and $\Pr\left(p_t^* \geq p_J\right) = 1 - F\left(p_J - \theta x_t\right)$. Similarly, the unconditional probability of not purchasing is

$$\begin{aligned}
\Pr\left(d_{0t} = 1\right) &= \Pr\left(p_t^* < p_1\right) \\
&\quad + \sum_{k=1}^{J-1} \Pr\left(d_{0t} = 1|p_k \leq p_t^* < p_{k+1}\right) \cdot \Pr\left(p_k \leq p_t^* < p_{k+1}\right) + \\
&\quad + \Pr\left(d_{0t} = 1|p_t^* \geq p_J\right) \cdot \Pr\left(p_t^* \geq p_J\right)
\end{aligned}$$

The conditional probability of choosing product j at time t is

$$\Pr\left(d_{jt} = 1|p_k \leq p_t^* < p_{k+1}\right) = \begin{cases} 0 & \text{if } p_j \geq p_{k+1} \\ \frac{\exp(\delta_{jt} + U(C_t) - T(I_t) + \rho V^e(I_t, p_t, a_t))}{1 + \sum_{l \in \Omega_t^*} \exp(\delta_{lt} + U(C_t) - T(I_t) + \rho V^e(I_t, p_t, a_t))} & \text{if } p_j < p_k \end{cases}$$

noting from the ordered sequence that Ω_t^* is the set of k products with the lowest prices (per pack), i.e., $\Omega_t^* = \{l |\, p_{lt} \leq p_t^*\}$.

We start by assuming that current prices are independent of past realizations of prices, i.e., $dF\left(p_t|p_{t-1}\right) = dF\left(p_t\right)$. Later we relax this assumption and allow for the possibility that prices follow an exogeneous first-order Markov process.

5 Estimation

The estimation of the parameters of the model relies on a nested fixed point algorithm where the solution of the dynamic problem is nested within the parameter search. The parameters of the model are estimated by maximizing the likelihood of observed choices. This likelihood is characterized by

$$\log L = \sum_{i,j,t} d_{ijt} \ln \Pr \left(d_{ijt} = 1 | s_t \right)$$

The algorithm consists of two loops. In the inner loop, given a set of parameters, we solve the dynamic problem and calculate the likelihood. The outer loop searches the parameter values that maximize the likelihood function. The numerical solution of the dynamic problem is obtained by value function iteration using a discrete approximation.

One of the challenges in implementing the algorithm is that we do not observe the products considered by the household, and hence we need to integrate over all potential sets. This strategy creates a considerable computational burden with a large number of products. To make the problem tractable, we make the following assumptions. First, products can be aggregated into ten brands: Tide, Xtra, Dynamo, Purex, All, Arm&Hammer, Era, Wisk, a Private Label, and a composite brand that includes all the other brands. This simplification drastically reduces the number of possible consideration sets. Second, each product can be assigned to one of three sizes: small, medium, and large. So, during each shopping trip households can only consider between 0 and 30 products.

We do not observe inventory or consumption decisions. The estimation of inventory holdings follows Hendel and Nevo (2006). For each household, we start with an initial guess for inventories and then calculate the inventory in each week from the observed purchases and the estimated consumption. To reduce the impact of the initial guess, the first 8 visits of each household are used to simulate the distribution of inventories but are not considered in the estimation of the likelihood.

To simplify the estimation procedure, we assume households are consuming detergent at a constant rate γ until they run out. We calculate the consumption rate as follows. We start by assuming the rate of consumption for each household is the weekly average purchases during the first 100 weeks. This initial guess therefore ignores the possibility of stock-outs. We then simulate the inventory in each week assuming that our guess for the consumption rate is the true consumption rate and calculate the number of weeks with no consumption in this simulation. We update the consumption rate correcting for the number of weeks with zero consumption (in the simulation). We repeat the previous procedure with the updated consumption rate until the difference between the consumption rate used to simulate inventory and the updated consumption rate is sufficiently small. We can recover the rate of consumption because we observe households over a long period of time. There is, however, some measurement error in the estimation of the consumption rate, but we believe that the measurement error is small as we observe several purchases for each household and the consumption rate is consumer-specific.

For the estimation we assume the nonprice observed attributes include only product display and feature ads. Hence, the parameters to be estimated are the price coefficient, the coefficients associated with product display and feature ads, the taste for a product, the parameterized functions for storage costs, and the cash-constraint.

We assume storage costs are quadratic: $T(I_t) = \theta_1 I_t + \theta_2 I_t^2$; and the utility of consumption is $u(C_t) = 0.1C_t$. The taste for a product is characterized by brand-size dummies.

The estimation procedure is performed using visits 9-41 of each household. We use purchases without store information to update inventories but we do not use those purchases in the estimation of the likelihood.

Given the heterogeneity in the price coefficients, income and consumption rate, without further aggregation, we would have to solve a dynamic programming problem for each household. To avoid that, we aggregate households into different types that vary by income, family size, birthyear, and consumption rate. The discount factor associated with the dynamic problem was set equal to 0.99.

6 Identification

If consumers were not cash-constrained, the identification of the price coefficient α_i and the brand-size dummies ξ_{jx} would be standard. Variation over time in prices and choices would identify the sensitivity to price. Household heterogeneity in price sensitivity would be captured by making the sensitivity to price a function of household income and size. Differences across households would therefore enable us to recover the heterogeneity in price sensitivity. Finally, the brand-size dummies would be identified from the variation in market shares across products. We can still use these variations to identify the price coefficient and the brand-size dummies if we have a sufficiently large number of observations for which the probability of considering all products (i.e. the household is not cash-constrained) is arbitrarily close to one. Essentially, we identify the parameters using only those observations, and then our model assumptions allow us to extrapolate these values for the remaining observations.

A common issue in the estimation of discrete choice models is the potential endogeneity problem that arises if prices are correlated with the unobserved variable ξ. We deal with potential endogeneity by (i) assuming that $\xi_{jxt} = \xi_{jx}$ and controlling it with fixed effects, (ii) controlling for displays and feature ads, and (iii) using weekly price data for each store.

As for the identification of cash constraints, we rely on exclusion restrictions. We assume some variables affect only consumers choices through the cash-constraint condition. So, the variation in these variables and in consumer choices identifies how these variables affect the products considered.

In particular, we assume that period of the month affects only the cash-constraint condition, because it is unlikely that period of the month affects the preferences for buying or consuming detergent. Hence, period of the month does not influence the choice conditional on the set of products that the household is able to consider, but it will affect the likelihood of considering a product in a given week. We can therefore identify the cash constraints through the variation created by period of the month in the likelihood of choosing each brand. For our identification of cash constraints, it is also important that we have variation in the ranking of products. Our arguments for identification of the cash constraint are therefore similar to the arguments used for identification of a standard selection model (Heckman 1978). In short, the identification of cash constraints is possible by (1) the availability of longitudinal data, (2) the observed heterogeneity on the likelihood of being cash constrained, and (3) the variation on the likelihood of being cash constrained within a household due to some observed variable that does not affect the value of purchasing and consuming. To build intuition for the identification suppose storage costs are the same for everyone, but rich people are never cash constrained. So, if we take two households who are identical in everything but income, we can identify the importance of cash-constraints by comparing the purchasing behavior in different weeks of the month. In particular, if the prices were constant and low income households had the largest cash constraint in the second half of each month (and were not cash constrained in the first half), then the difference between the second and first half of each month on the differences of purchases between high and low income households would be explained by the presence of cash constraints.

The identification of consumption behavior is critical for the identification of the utility and storage costs. If inventory and consumption were observed, the identification of functions of these variables would follow standard arguments. We do not observe inventory and consumption and thus we need to be able to identify consumption behavior.

The possibility of purchasing containers of different sizes is one of the key elements for the identification of both storage costs and liquidity constraints. The purchase of a small container can be explained by either large storage costs or large liquidity constraints. It is therefore important to distinguish between storage costs and liquidity constraints as both can yield the purchase of small containers. The identification of these two functions of the model is possible because of the exclusion restrictions that we make regarding storage costs and liquidity constraints. In particular, we assume they depend on different household demographics–the cash constraint depends on income, whereas storage costs are similar across households–and the cash constraint depends on the period of the month, whereas the storage technology does not vary over time. Moreover, storage costs depend on current inventory, whereas the cash constraint is affected by actual prices. So, we expect storage costs to be high if purchases of small containers are higher when the time elapsed since the last purchase is shorter.[15] Finally, as we have already pointed out, if some households are not cash-constrained, the purchase of small containers is only explained by storage costs and therefore we can use the usual arguments for identification of storage costs.

7 Results

7.1 Parameter Estimates

In table 6 we report the parameters for two specifications of the previously described dynamic demand model with cash constraints. The results reveal that cash-constraints are statistical and economically significant. Also, they depend on household income and the period of the month when the shopping trip takes place. Cash-constraints are higher in the third week of the month, when there is a high probability of considering none or only a small number of products. Most of the wages (or government benefits) are paid in the beginning of the month and therefore the available cash tends to fall over the month, which can justify our results. Interestingly, cash-constraints seem to be lower in the second week of the month rather than in the first week of the month. One potential explanation is a potential lag between the time when households receive their monthly allowance and the time of the first shopping trip after receiving it. Income has an important effect on cash-constraints. Households with income above \$60000 have indeed very small, if any, cash constraints, whereas households with income below \$15000 have considerably high cash-constraints. In order to illustrate the importance of cash-constraints and the impact of income on them, figures 20 and 21 (panel b) plot the distribution of the number of products considered by a household during a shopping using the estimates of the second specification in table 6. Figure 20 considers all households, whereas figure 21 considers only low-income households. Figure 20 reveals that the probability of households not having enough cash to consider all the available products is nearly 20%. There is, however, a large heterogeneity on the level of cash constraints with respect to household income. As figure 21 shows, the positive values on the left tail of the distribution of the number of considered products are essentially due to households with low-income. In fact, the probability of low-income households not having cash to consider all products is larger than 75%. Moreover, the probability of these households considering at most 20 products is larger than 20%, which means that it is unlikely that they will purchase large sizes. The previous points are reinforced by figure 22, which displays the average number of products considered during a shopping trip for all households and for households with low income. Households, on average, have cash to consider around 96% of the available products, but households with income below \$15000 consider only 84% of the available products and have a probability higher than 15% of considering at most 2 brands in large containers (figure 23).

[15] The same argument is valid with respect to the quantity purchased during the last purchase occasion.

The storage cost and the coefficients for price, product display and feature ads are significant and have the expected sign. The estimates of storage costs are, however, small, which is partly driven by the presence of households with very large estimated inventory holdings and our specifications do not allow for heterogeneous storage costs.

7.2 Model Fit and Comparison with Alternative Models

In this subsection I compare the propose model with alternative models. The objective of this comparison is twofold. The first goal is to evaluate the biases created by ignoring cash constraints, forward looking behavior, or demand accumulation. The second is to evaluate the fit of the model of proposed model. Table 7 reports the estimates for alternative choice models. Columns 1 and 2 show the results for a static model without cash constraints; column 3 reports the estimates for a static model with cash constraints.

The results in column 3 reveal that ignoring forward looking behavior and demand accumulation largely overestimates cash constraints. This overestimation is well illustrated by figures 20 to 24, which show that a static model predicts that consumers consider only nearly 10% of the available products and have a probability close to 90% of considering 0 products during a shopping trip. Cash constraints and price elasticity are the only mechanisms that allow the model to capture the households' choice of not buying a product during a shopping trip in a static model without storage costs. As we observe a large number of shopping trips without purchases, the model rationalizes this with large cash constraints. In contrast, a dynamic model with demand accumulation can explain the decision of not purchasing with the postponement to the next trip or large storage costs, thereby lowering the impact of cash constraints. Interestingly, the effect of income and week of the month on the cash constraint is very small in a static model. The reason is that we observe decisions of not purchasing for all income groups and across all weeks, so the cash constraint needs to be always high so that it can explain the large share of shopping trips without purchases.

8 Counterfactuals

We use the estimates of our model to evaluate the welfare implications of cash-constraints. In particular, we calculate how expenditures, consumption and choice of container-size would be different without a cash constraint. To do that, we simulate households choice with and without a cash constraint using the estimates of the model. For the situation without a cash constraint we assume that all parameters are similar to the situation with a cash constraint but the household considers all available products with probability 1 during all shopping trips. In our analysis we give particular attention to the effects for low income households, because these households are more susceptible of being cash constrained and therefore removing the cash constraint can be particularly important for them.

9 References

Bronnenberg, B., M. Kruger and C.Mela (2008), Database paper: The IRI marketing data set, *Marketing Science 27, 745-748*

Chu, Arlene and Aviv Nevo (2013), The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession, *Northwestern manuscript*

Cohen, Andrew (2008), Package Size and Price Discrimination in the Paper Towel Market, *International Journal of Industrial Organization 26, 512-516*

Donna, Javier, and Jose-Antonio Espin-Sanchez (2014), The Illiquidity of Water Markets, *Northwestern manuscript*

Dreze, X., S.Hoch and M.Purk (1994), Shelf Management and Space Elasticity, *Journal of Retailing 70, 301-326*

Gilbride, Timothy J. and Greg M. Allenby (2004), A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules, *Marketing Science 23, 391-406*

Hastings, Justine, and Jesse Shapiro (2013), Fungibility and Consumer Choice: Evidence from Commodity Price Shocks, *The Quarterly Journal of Economics 128, 1449-1498*

Heckman, James. (1979), Sample selection bias as a specification error, *Econometrica 47, 153–161*

Hendel, Igal, Alessandro Lizzeri, and Nikita Roketskiy (2014), Nonlinear Pricing of Storable Goods, *American Economic Journal: Microeconomics 6, 1-34*

Hendel, I. and A. Nevo (2006), Measuring the Implications of Sales and Consumer Inventory Behavior, *Econometrica 74, 1637 - 1673*

Hoch, S. and J.Deighton (1989), Managing What Consumers Learn From Experience, *Journal of Marketing 53, 1-20*

Huffman, David and Matias Barenstein (2005), A Monthly Struggle for Self-Control? Hyperbolic Discounting, Mental Accounting, and the Fall in Expenditure Between Paydays, *IZA Discussion Paper, No. 1430*

Jappelli, Tulio, and Luigi Pistafferi (2010), The Consumption Response to Income Changes, *Annual Review of Economics 2, 479-506*

Kruger, M. and D. Pagni (2008), IRI Academic Data Set Description 1.403, *Chicago: Information Resources Incorporated*

McManus, Brian (2007), Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee, *Rand Journal of Economics 38,* 512–532

Nevo, A. (2001), Market Power in the Ready-to-Eat Cereal Industry, *Econometrica 69, 307 - 342*

Rust, J. (1987), Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher, *Econometrica 55, 999-1033*

Seiler, S. (2013), The Impact of Search Costs on Consumer Behavior: A Dynamic Approach *Quantitative Marketing and Economics 11, 155-203*

Shapiro, Jesse (2005), Is There a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle, *Journal of Public Economics 89, 303-325*

Skinner, Jonathan (1987), A Superior Measure of Consumption from the Panel Study of Income Dynamics, *Economic Letters 23, 213-287*

Stephens, Melvin (2001), The Long-Run Consumption Effects of Earning Shocks, *Review of Economics and Statistics 83, 28-36*

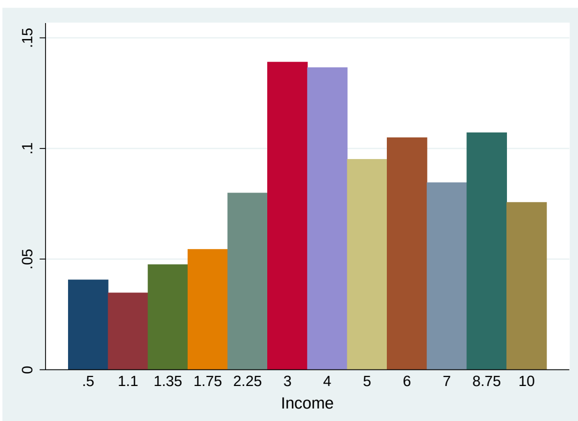
Stephens, Melvin (2004), Job Loss Expectations, Realizations, and Household Consumption Behavior, *Review of Economics and Statistics 86, 253-269*

Stephens, Melvin (2006), Paycheque Receipt and the Timing of Consumption, *The Economic Journal 116, 680-701*

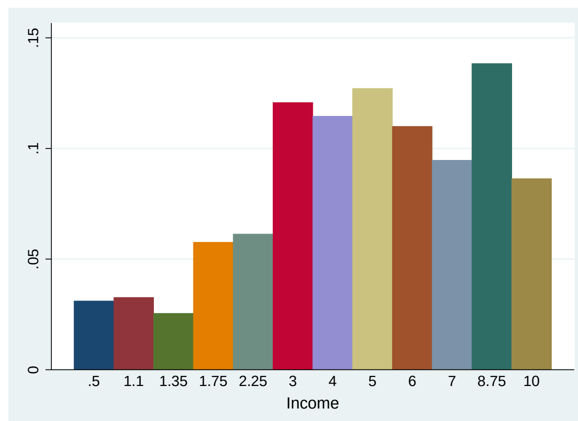
Tarozzi, Alessandro, Aprajit Mahajan, Brian Blackburn, Dan Kopf, Lakshmi Krishnan, and Joanne Yoong (2014), Micro-Loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India, *American Economic Review 104, 1909-1941*

10 Figures

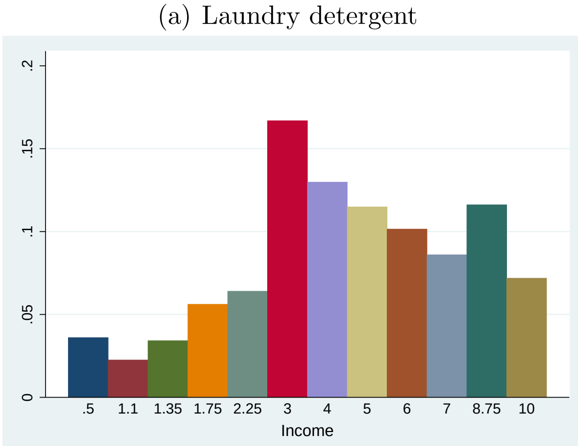
Figure 1: Income Distribution



(a) Laundry detergent



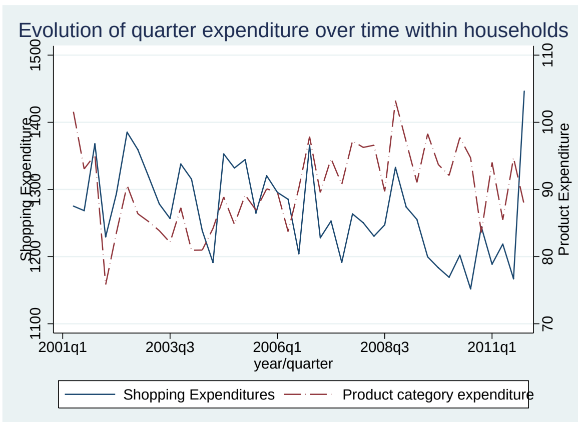
(b) Shampoo



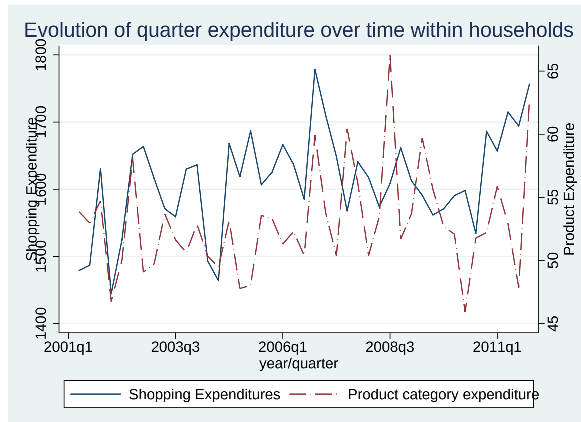
(c) Toothpaste

Note: Each panel reports the percentage of households in each income level for the sample of purchases in the designated product category. Panel (a) reports the distribution for households in the sample of purchases of laundry detergent, panel (b) reports the distribution for households in the sample of purchases of shampoo, and panel (c) report the distribution for households in the sample of purchases of toothpaste.

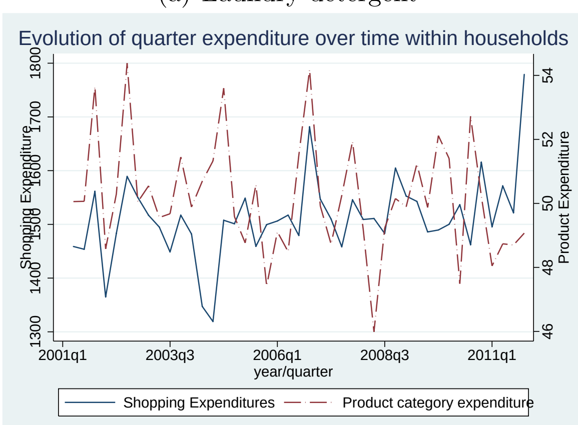
Figure 2: Evolution of quarter expenditures within households



(a) Laundry detergent



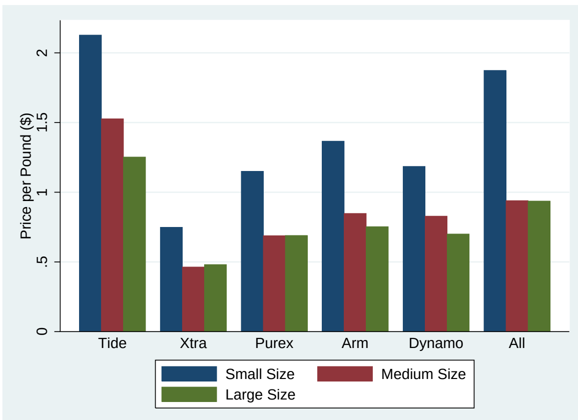
(b) Shampoo



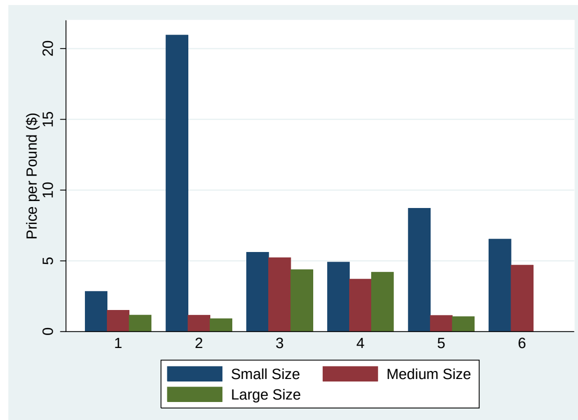
(c) Toothpaste

Note: Each panel shows the evolution of total quarter expenditures (in grocery stores) and quarter expenditures of a specific product category. The values reported are the sum of the residual and constant of regressing shopping trip expenditures on household fixed effects. The values are then aggregated by quarter for each household. The three panels show outcomes for a different product category and using the sample of households with purchases in that product category (each panel therefore considers a different sample)

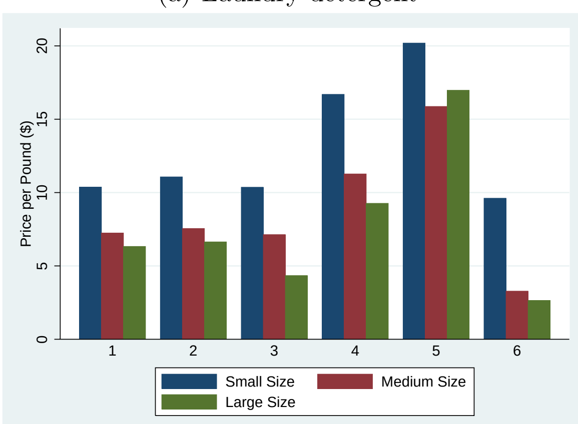
Figure 3: Price Per Pound



(a) Laundry detergent



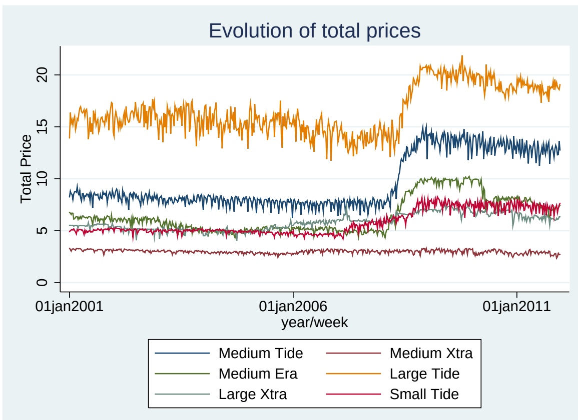
(b) Shampoo



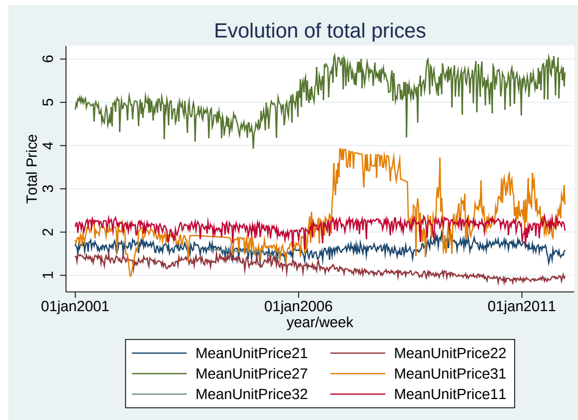
(c) Toothpaste

Note: Each panel shows the price per pound of different brand sizes for the designated product category.

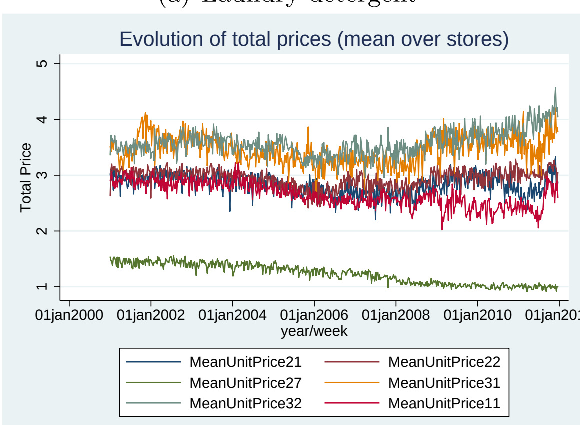
Figure 4: Evolution of prices



(a) Laundry detergent



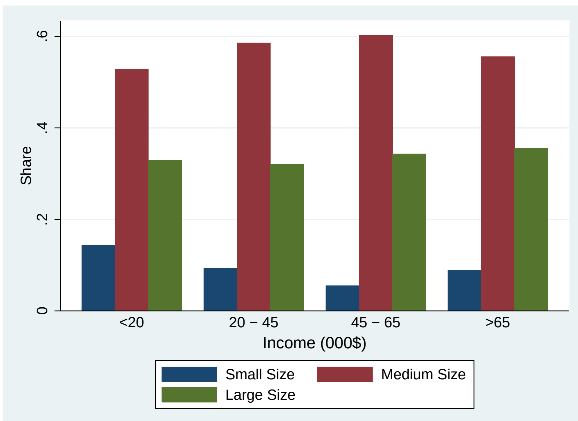
(b) Shampoo



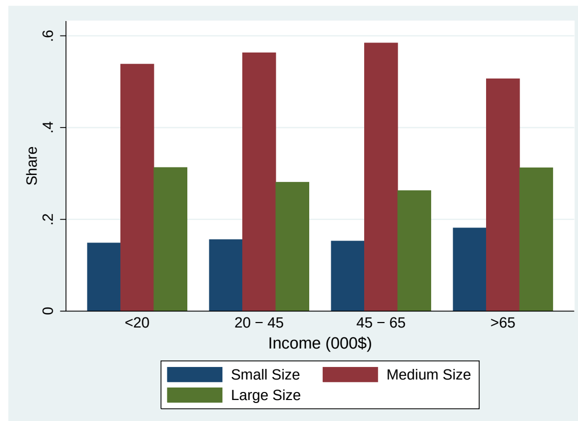
(c) Toothpaste

Note: Each panel presents for the designated product category the price per pound of 6 products averaged across all stores over 11 years beginning January 2001.

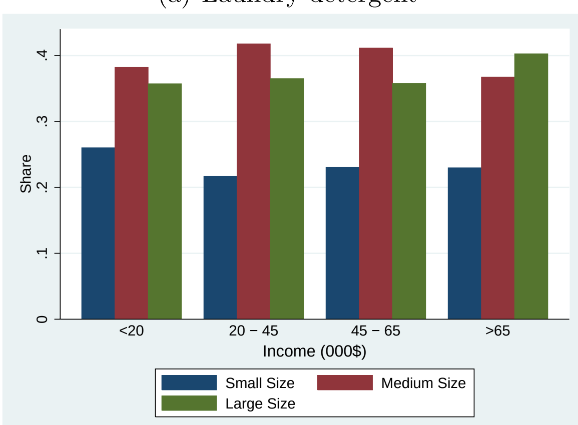
Figure 5: Share of purchases of each container's size stratified by household income



(a) Laundry detergent



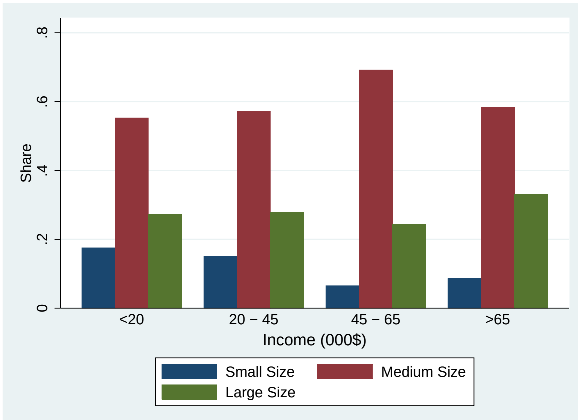
(b) Shampoo



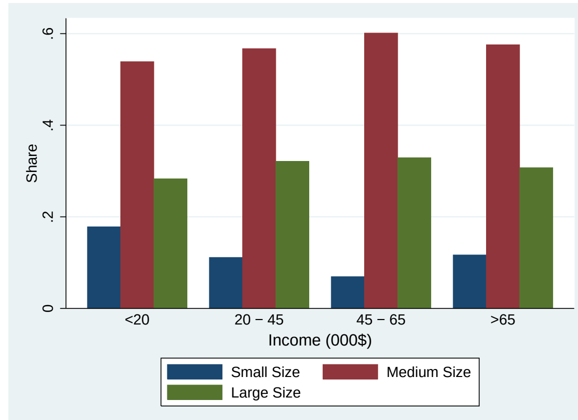
(c) Toothpaste

Note: Each panel shows share of purchases of each container's size stratified by household income for the designated product category.

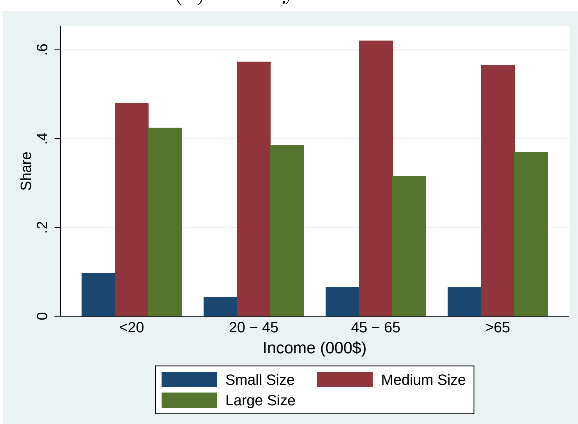
Figure 6: Share of purchases of each container's size stratified by household income and size - Laundry detergent



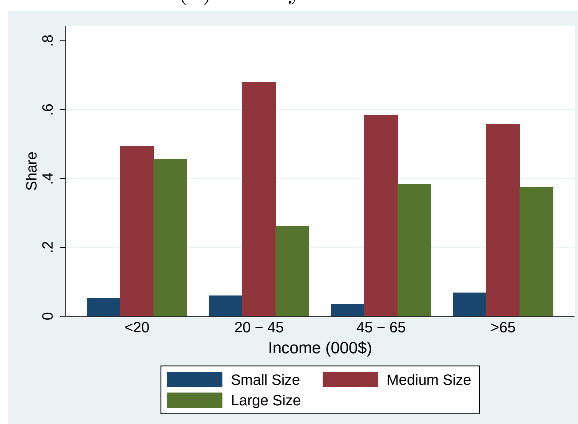
(a) Family Size = 1



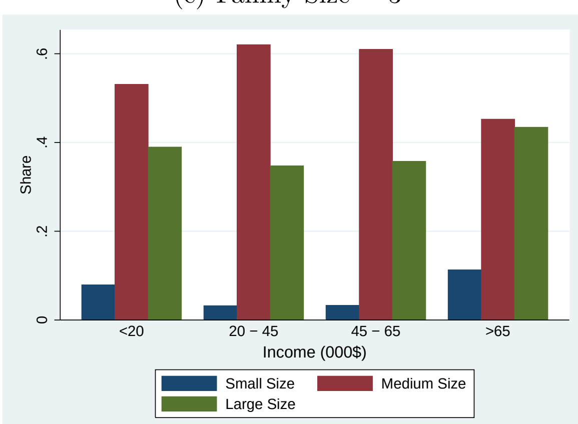
(b) Family Size = 2



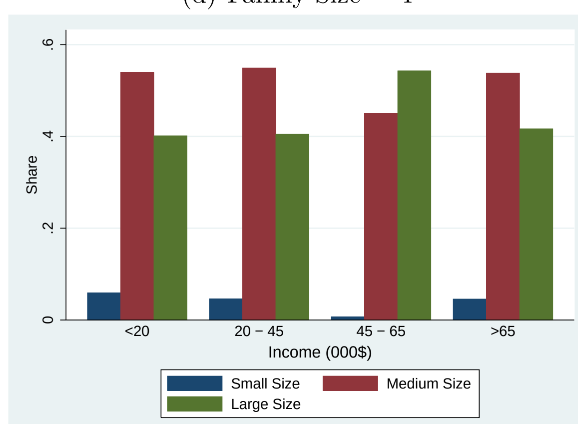
(c) Family Size = 3



(d) Family Size = 4



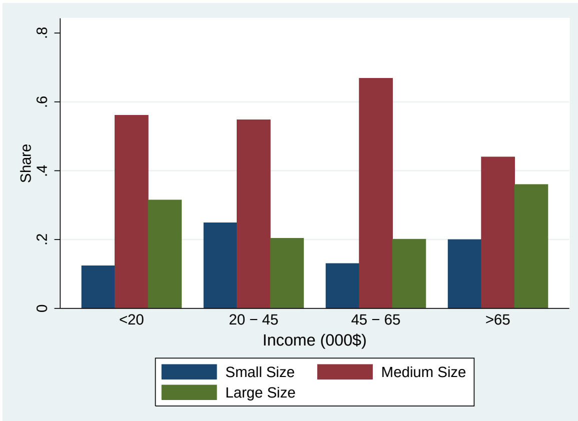
(e) Family Size = 5



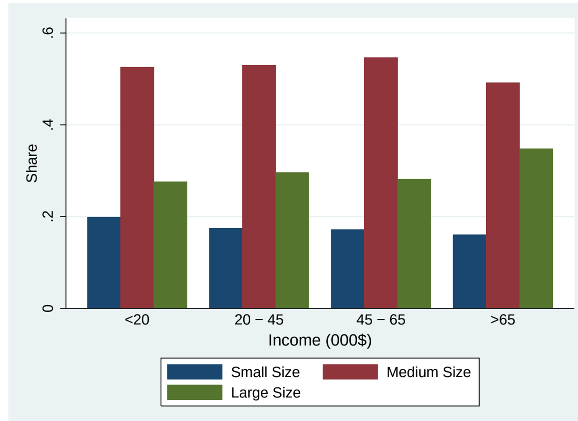
(f) Family Size > 5

Note: Each panel shows share of purchases of each container's size stratified by household income and size for laundry detergent.

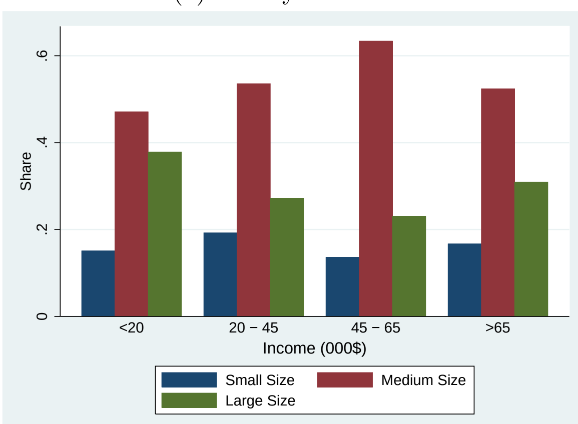
Figure 7: Share of purchases of each container's size stratified by household income and size - Shampoo



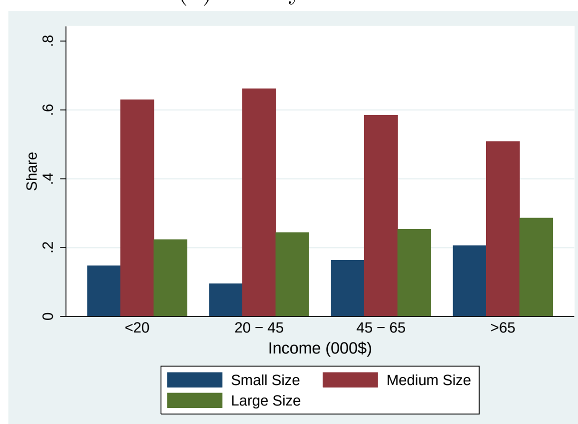
(a) Family Size = 1



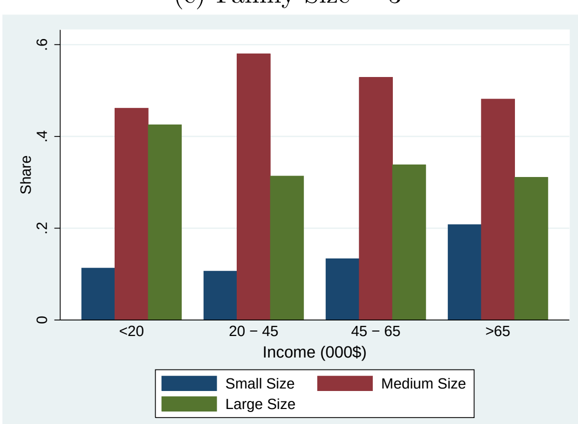
(b) Family Size = 2



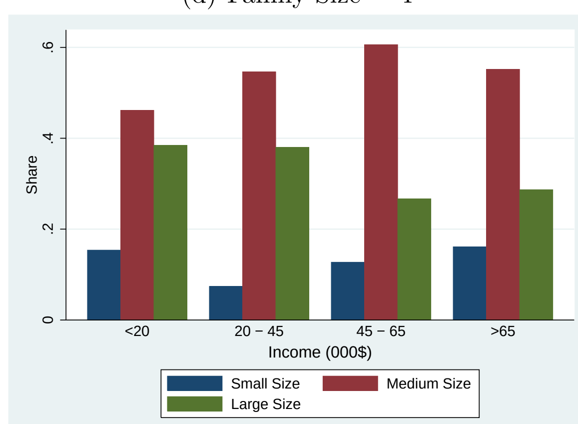
(c) Family Size = 3



(d) Family Size = 4



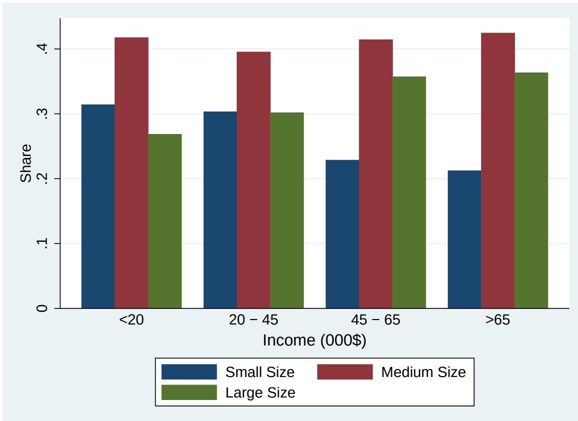
(e) Family Size = 5



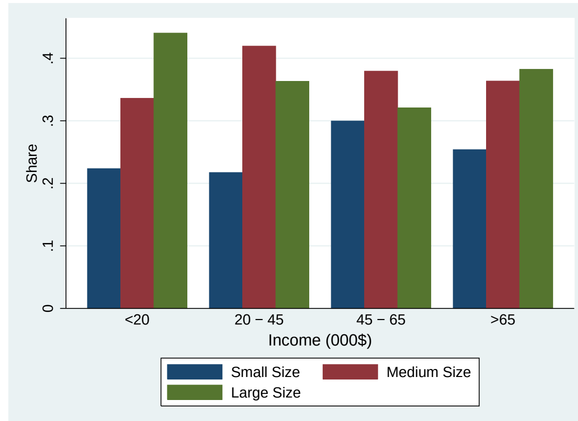
(f) Family Size > 5

Note: Each panel shows share of purchases of each container's size stratified by household income and size for shampoo.

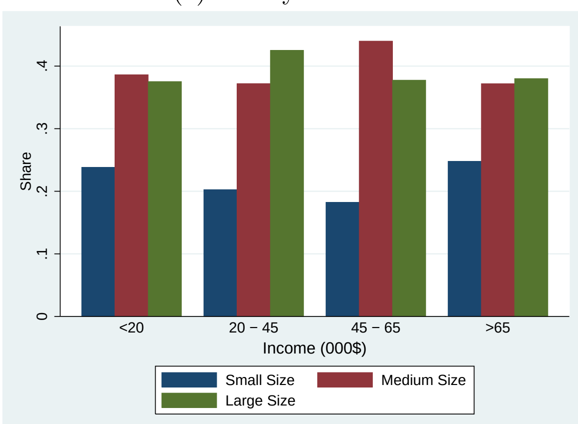
Figure 8: Share of purchases of each container's size stratified by household income and size - Toothpaste



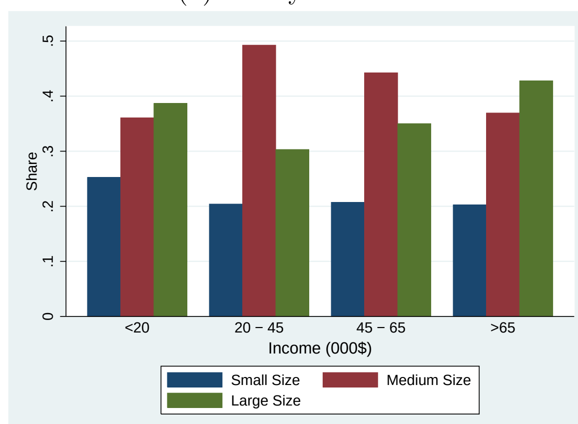
(a) Family Size = 1



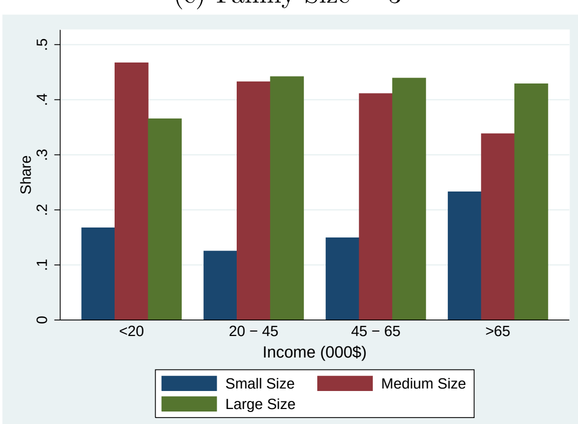
(b) Family Size = 2



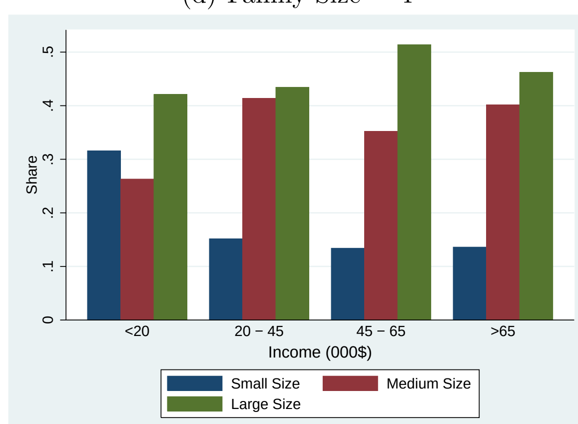
(c) Family Size = 3



(d) Family Size = 4



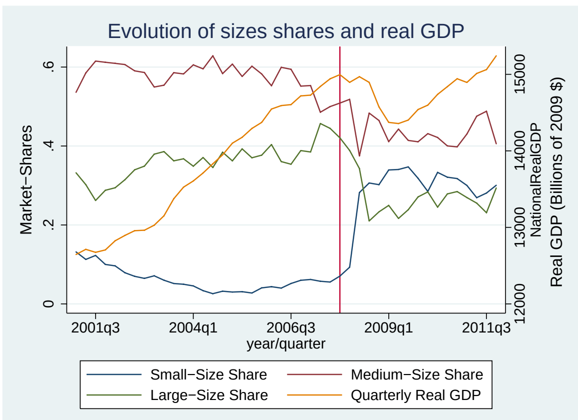
(e) Family Size = 5



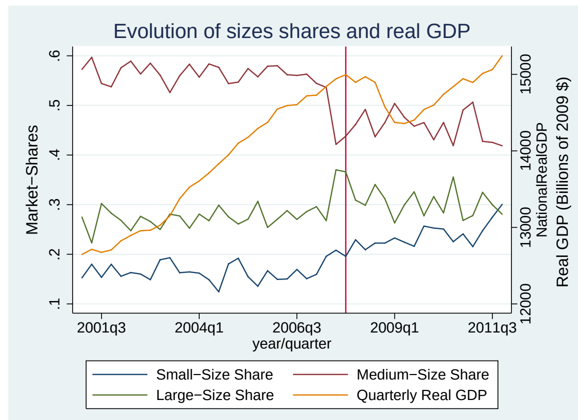
(f) Family Size > 5

Note: Each panel shows share of purchases of each container's size stratified by household income and size for toothpaste.

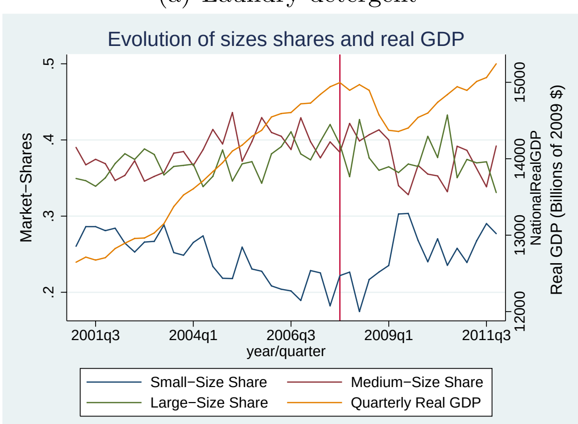
Figure 9: Quarterly evolution real GDP and sizes share



(a) Laundry detergent

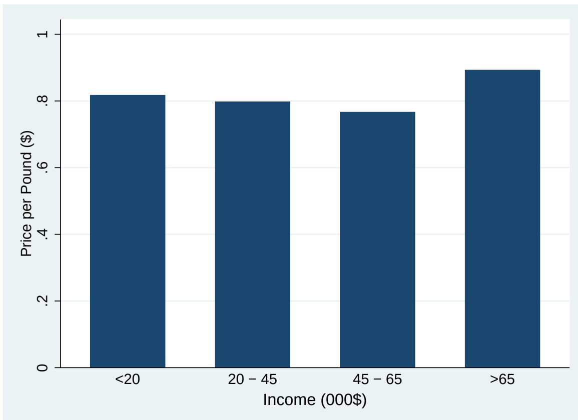


(b) Shampoo

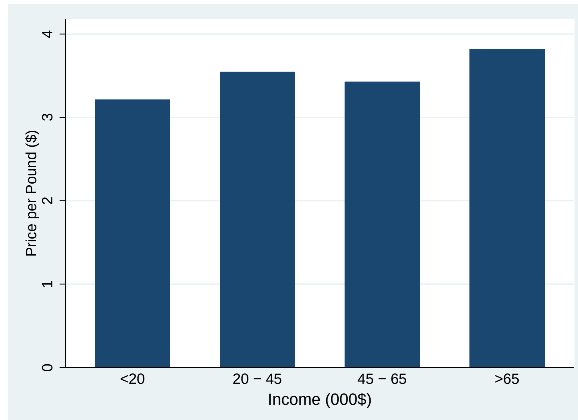


(c) Toothpaste

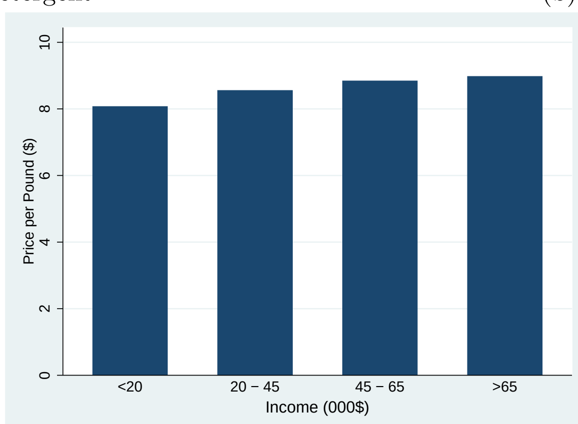
Figure 10: Price by Income



(a) Laundry detergent



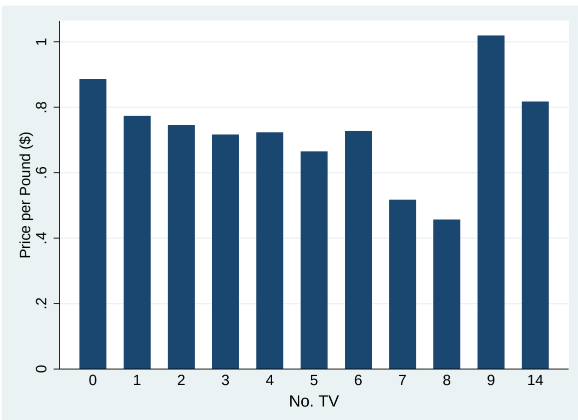
(b) Shampoo



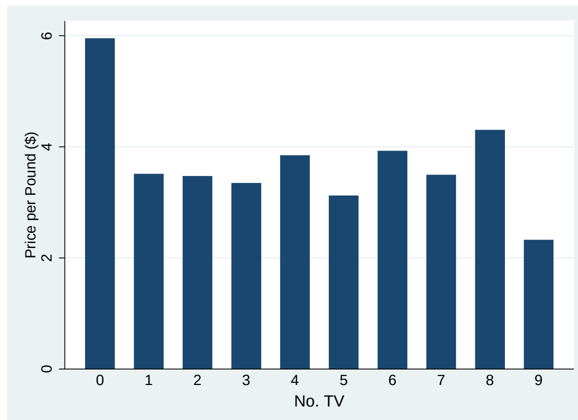
(c) Toothpaste

Note: Each panel reports the transaction price per pound for each income group (i.e., transaction prices are averaged over all purchases of households in the designated income group). Income is the annual combined pre-tax income of the heads of the household and is in thousands of dollars. Panel (a) reports the values for laundry detergent, panel (b) for shampoo, and panel (c) for toothpaste.

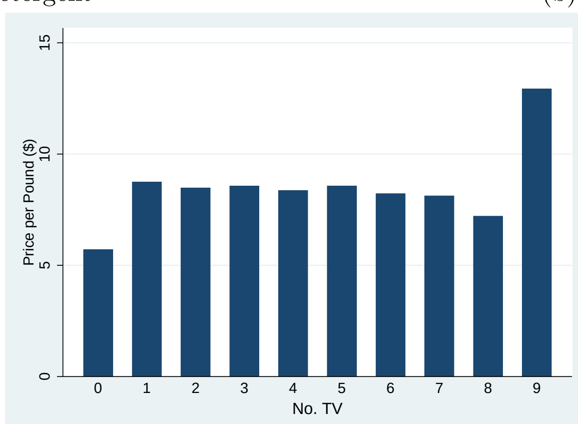
Figure 11: Price by No. TVs



(a) Laundry detergent

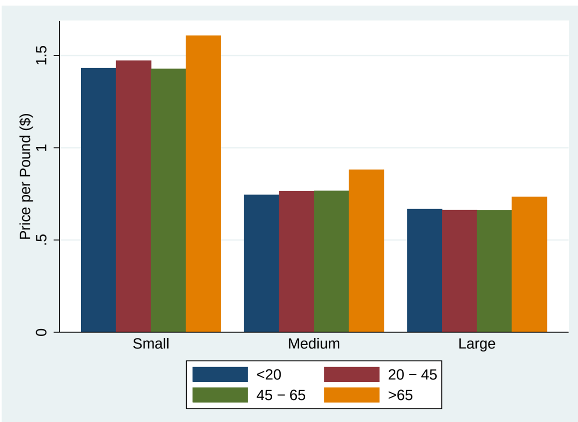


(b) Shampoo

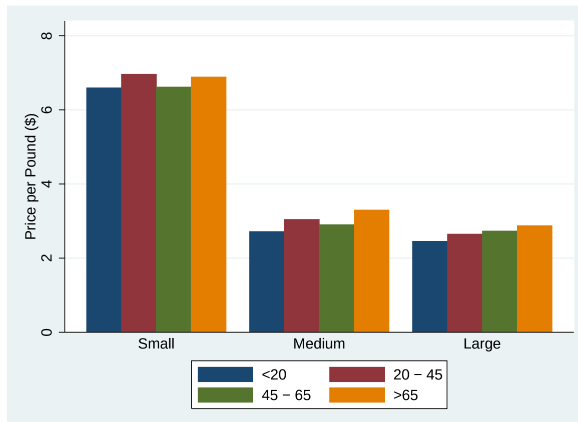


(c) Toothpaste

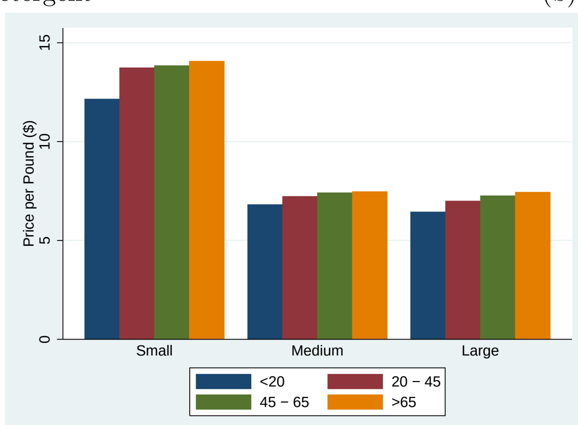
Figure 12: Price by Income stratified by size purchased



(a) Laundry detergent

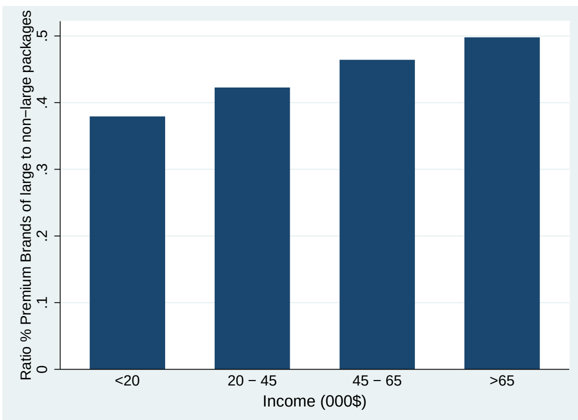


(b) Shampoo

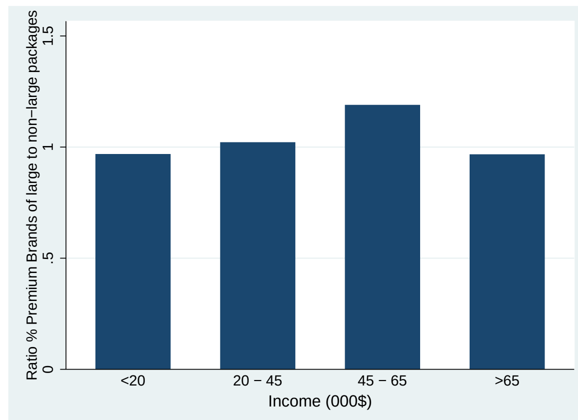


(c) Toothpaste

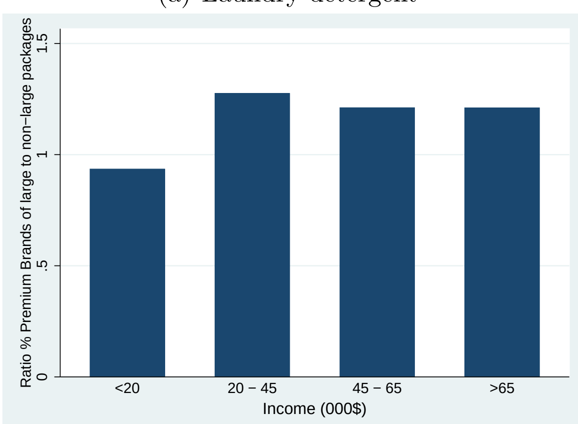
Figure 13: Ratio of Share of premium brands during purchases of large packages to Share of premium brands during purchases of non-large packages by household income (purchases)



(a) Laundry detergent

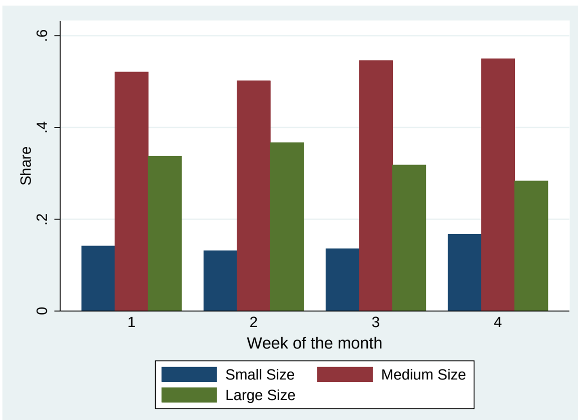


(b) Shampoo

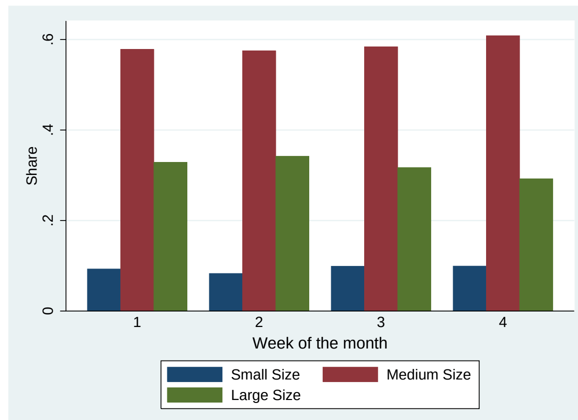


(c) Toothpaste

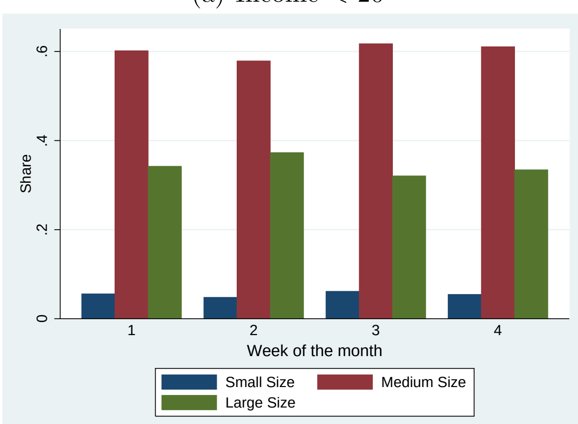
Figure 14: Share of each size by week of the month stratified by Income - Laundry detergent



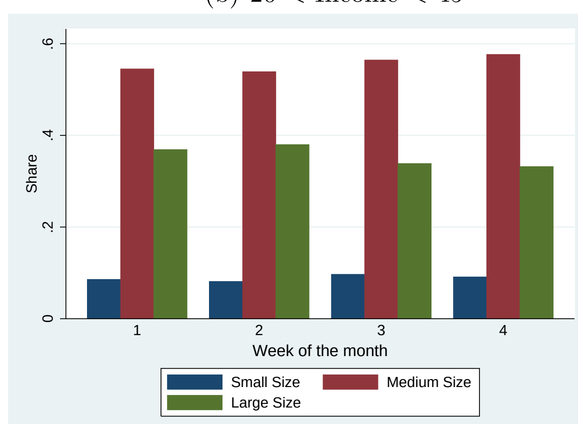
(a) Income < 20



(b) $20 <$ Income < 45

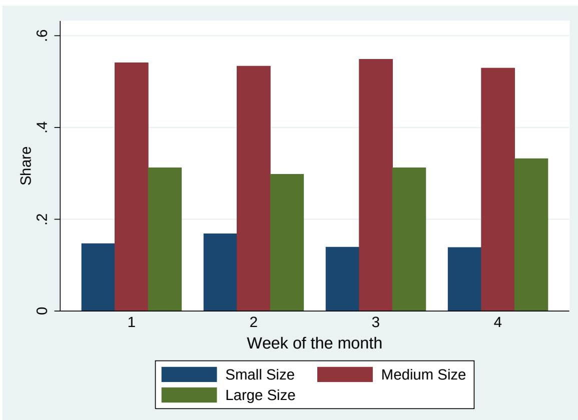


(c) $45 <$ Income < 65

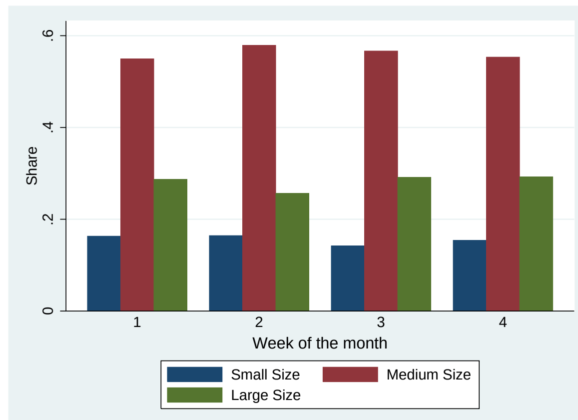


(d) Income > 65

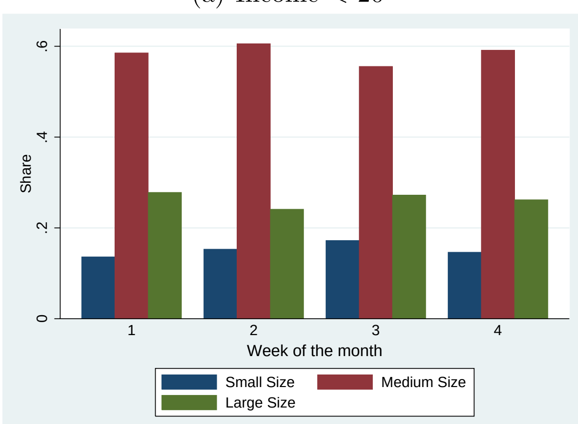
Figure 15: Share of each size by week of the month stratified by Income - Shampoo



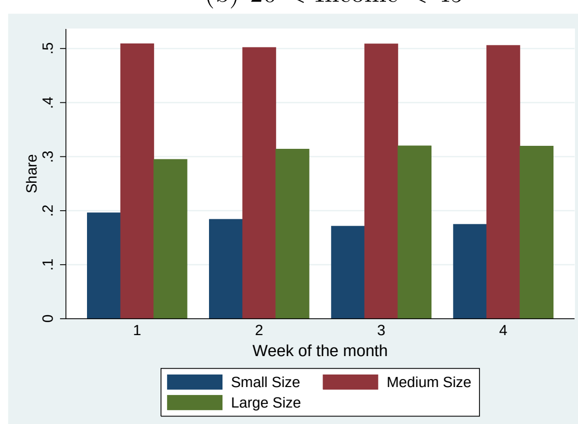
(a) Income < 20



(b) 20 < Income < 45

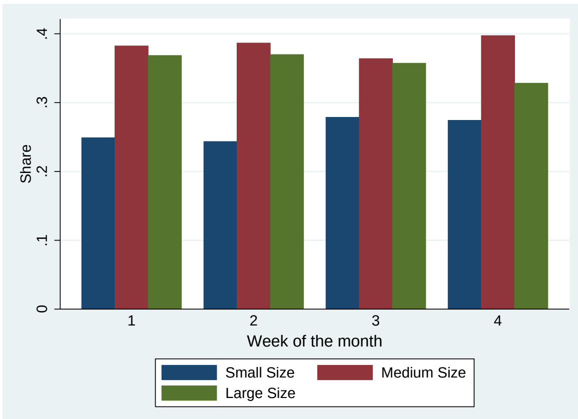


(c) 45 < Income < 65

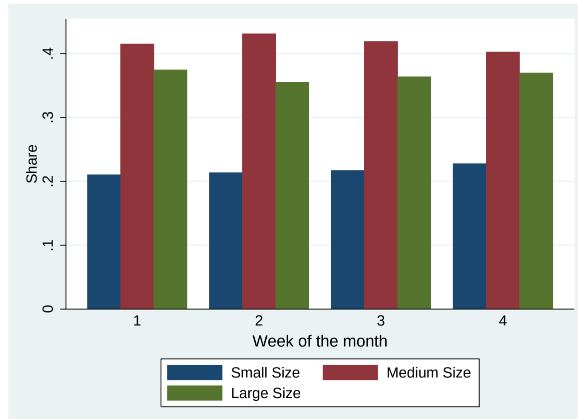


(d) Income > 65

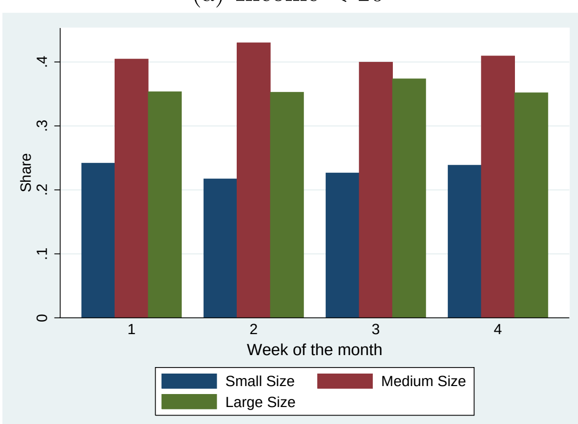
Figure 16: Share of each size by week of the month stratified by Income - Toothpaste



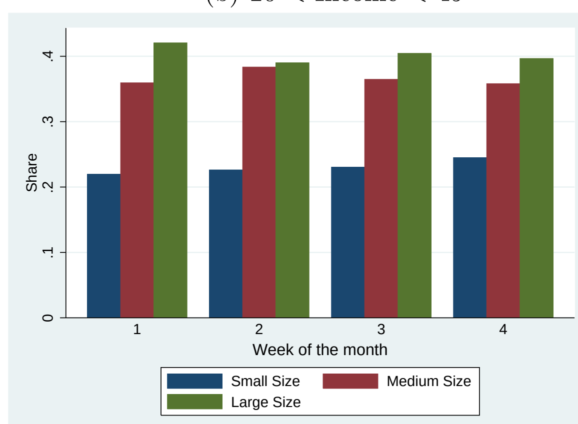
(a) Income < 20



(b) $20 <$ Income < 45

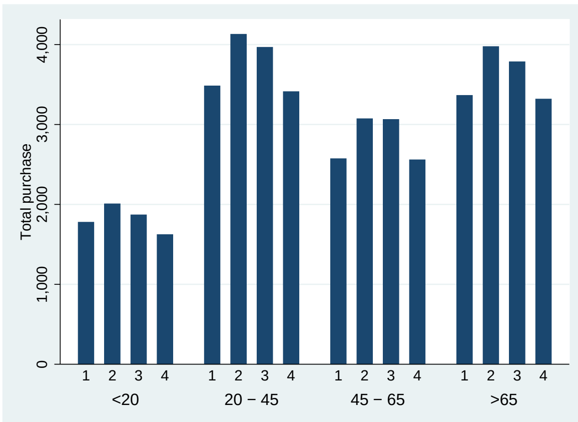


(c) $45 <$ Income < 65

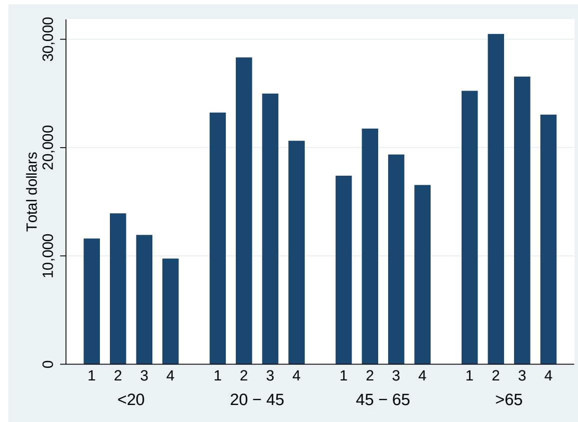


(d) Income > 65

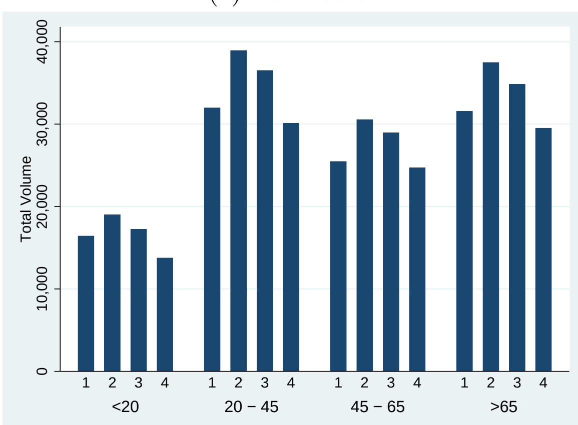
Figure 17: Outcomes stratified by income and week of the month - Laundry detergent



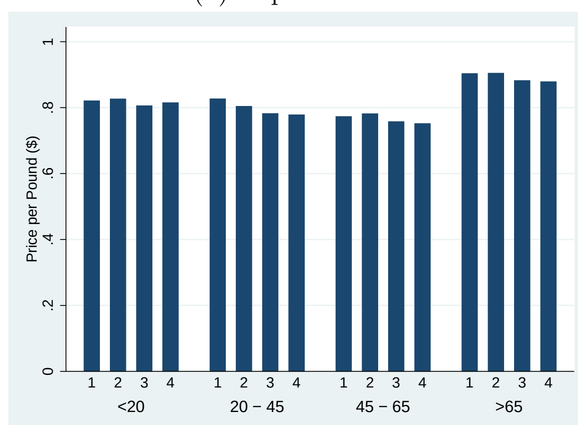
(a) Purchases



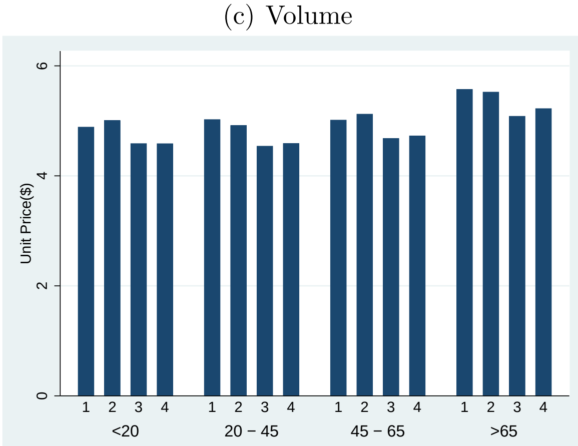
(b) Expenditures



(c) Volume

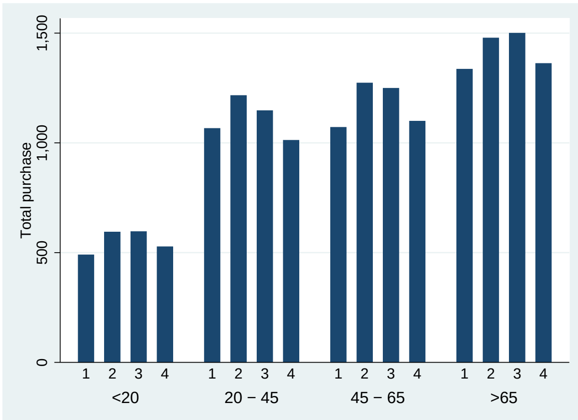


(d) Price Per Pound

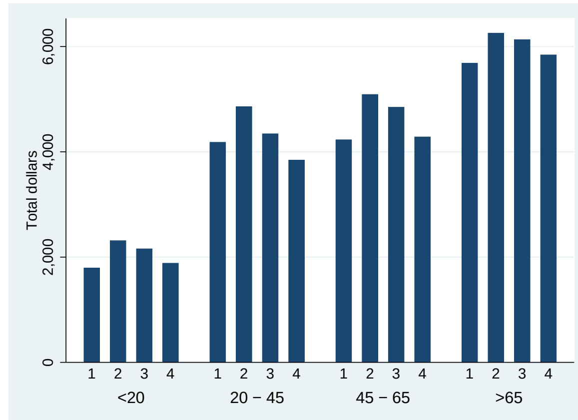


(e) Unit Price

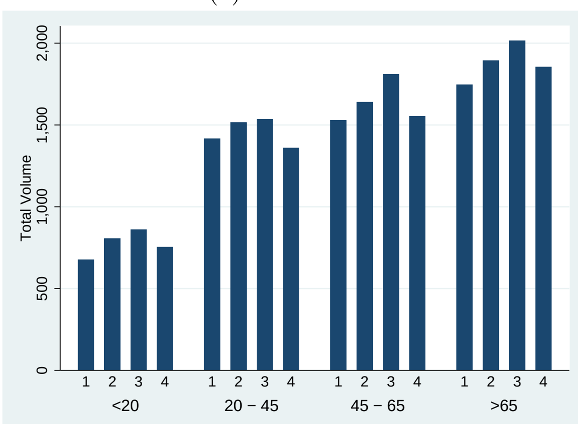
Figure 18: Outcomes stratified by income and week of the month - Shampoo



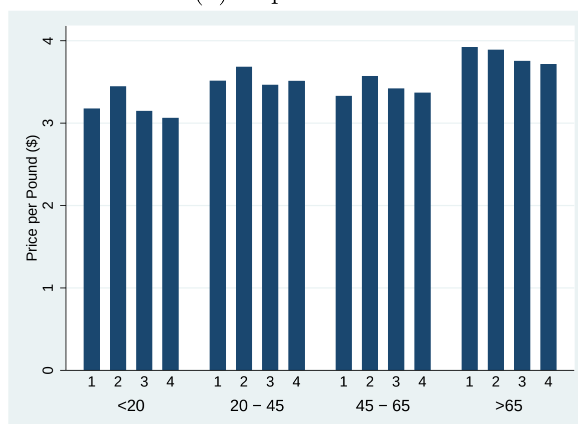
(a) Purchases



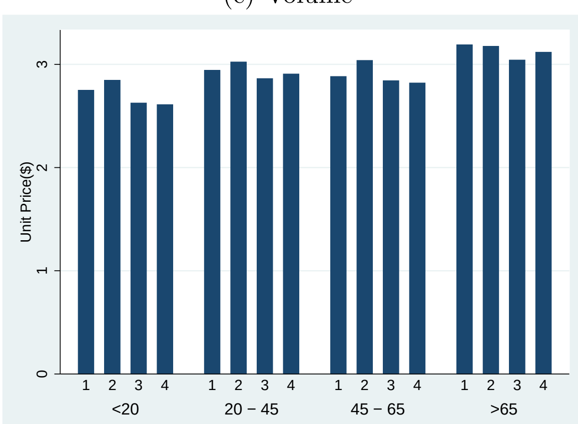
(b) Expenditures



(c) Volume

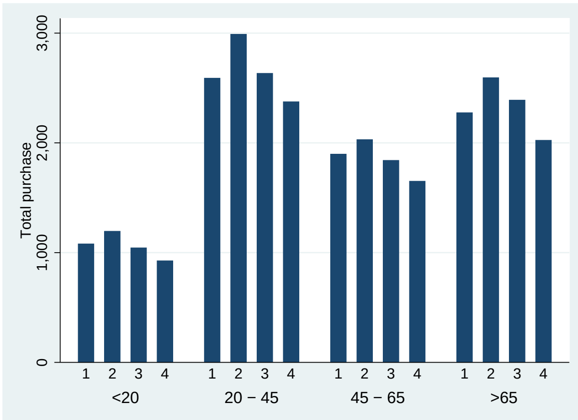


(d) Price Per Pound

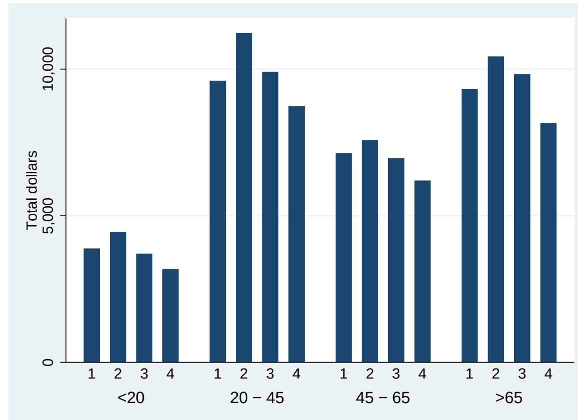


(e) Unit Price

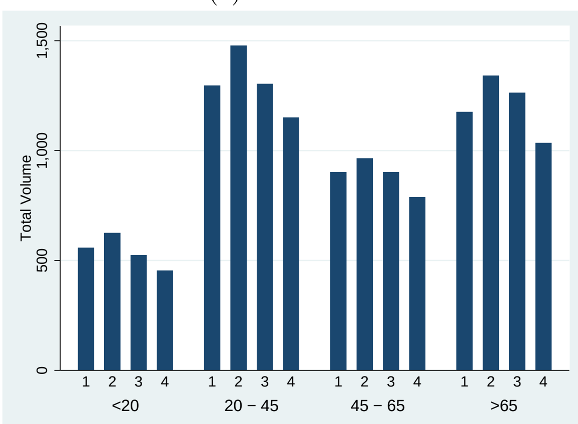
Figure 19: Outcomes stratified by income and week of the month - Toothpaste



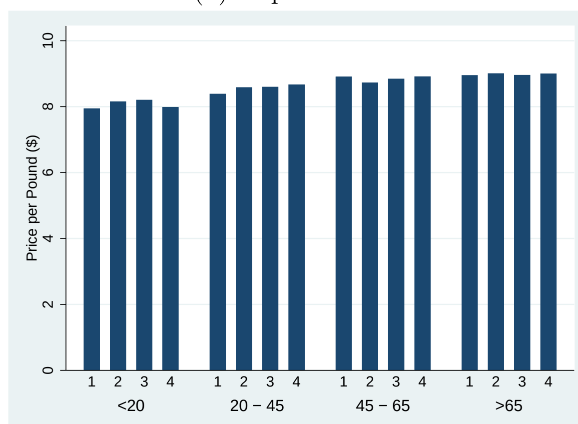
(a) Purchases



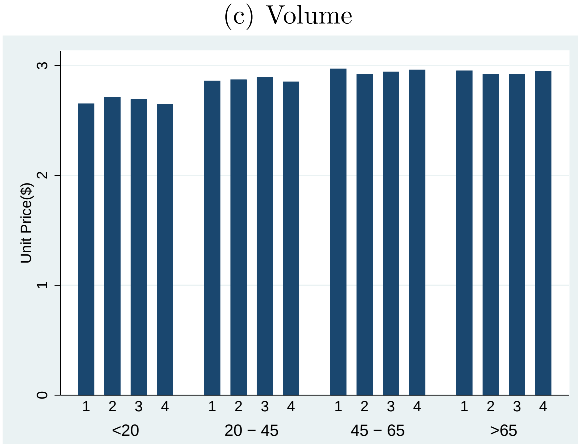
(b) Expenditures



(c) Volume

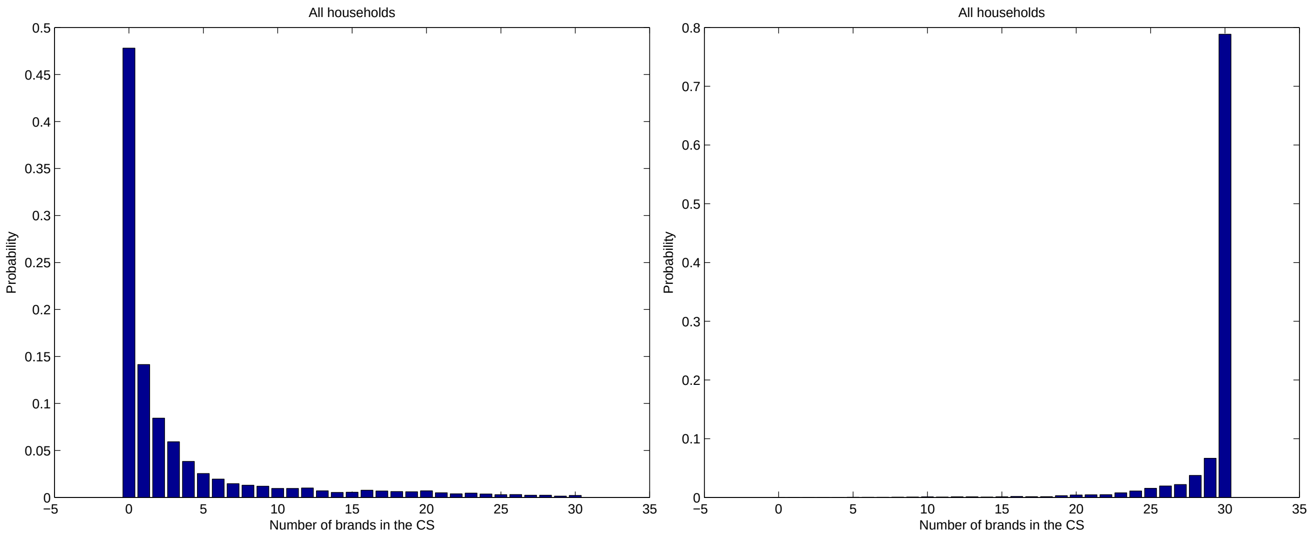


(d) Price Per Pound



(e) Unit Price

Figure 20: Distribution of number of considered products for all households (observation is a shopping trip)

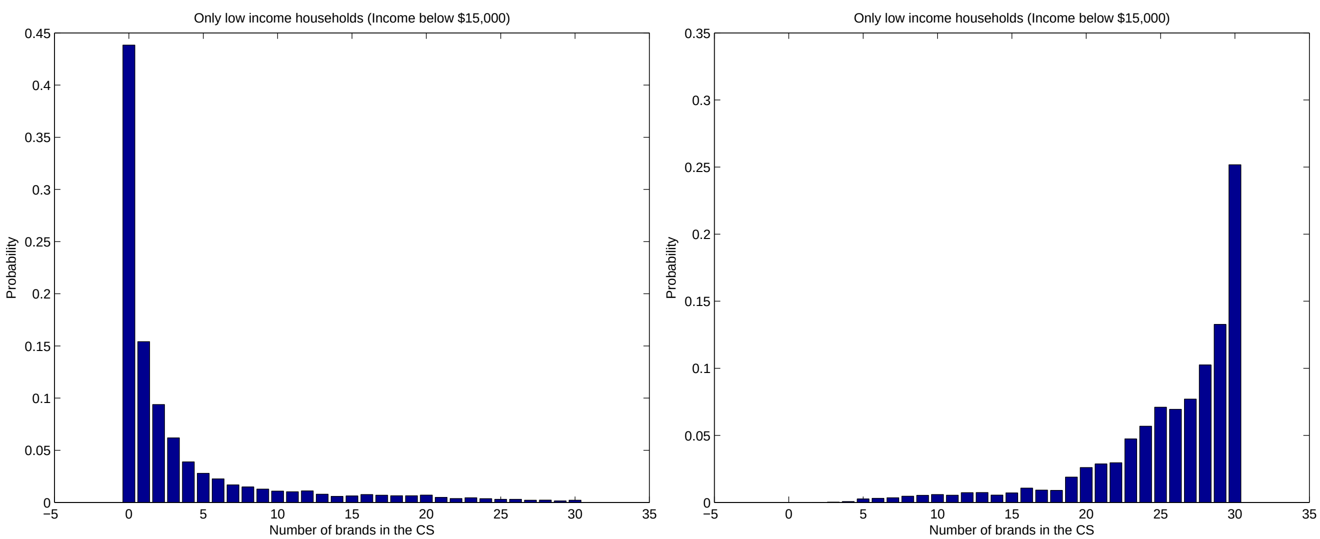


(a) Static Model

(b) Dynamic Model

Note: Each panel reports the probability distribution of the number of products considered during each shopping trip in the designated model. An observation is a shopping trip and the sample includes all households. Panel (a) reports the distribution using the estimates of the static model with a cash constraint reported in the third column of table 7; panel (b) reports the distribution using the estimates of the dynamic model with a cash constraint reported in the second column of table 6.

Figure 21: Distribution of number of considered products for low income households (observation is a shopping trip)

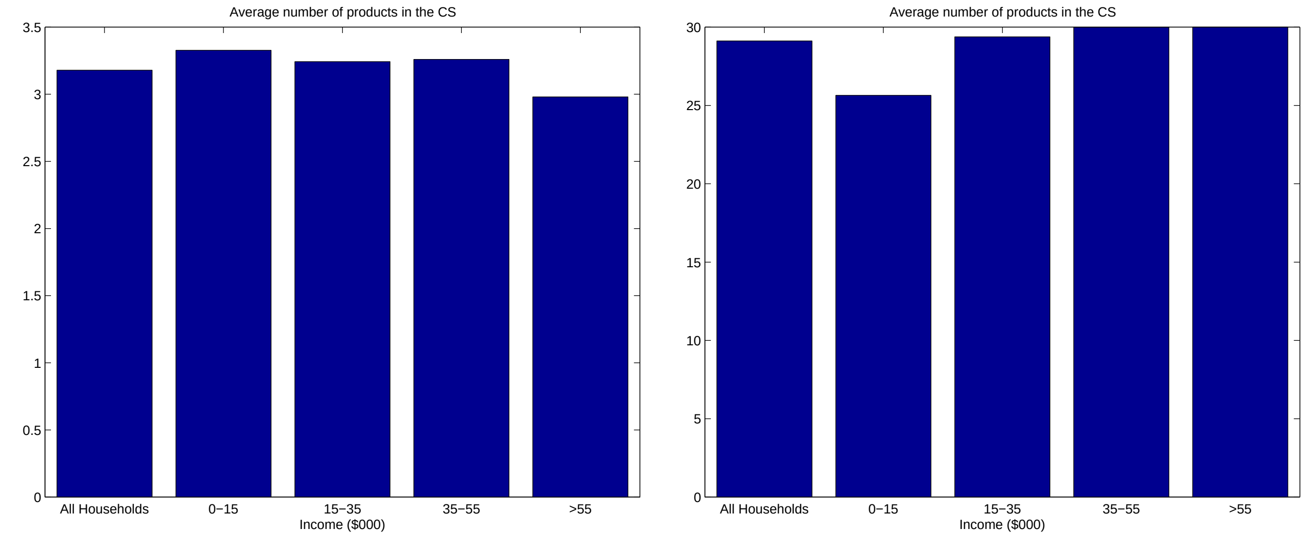


(a) Static Model

(b) Dynamic Model

Note: Each panel reports the probability distribution of the number of products considered during each shopping trip in the designated model. An observation is a shopping trip and the sample includes only households with annual income lower than $15000. Panel (a) reports the distribution using the estimates of the static model with a cash constraint reported in the third column of table 7; panel (b) reports the distribution using the estimates of the dynamic model with a cash constraint reported in the second column of table 6.

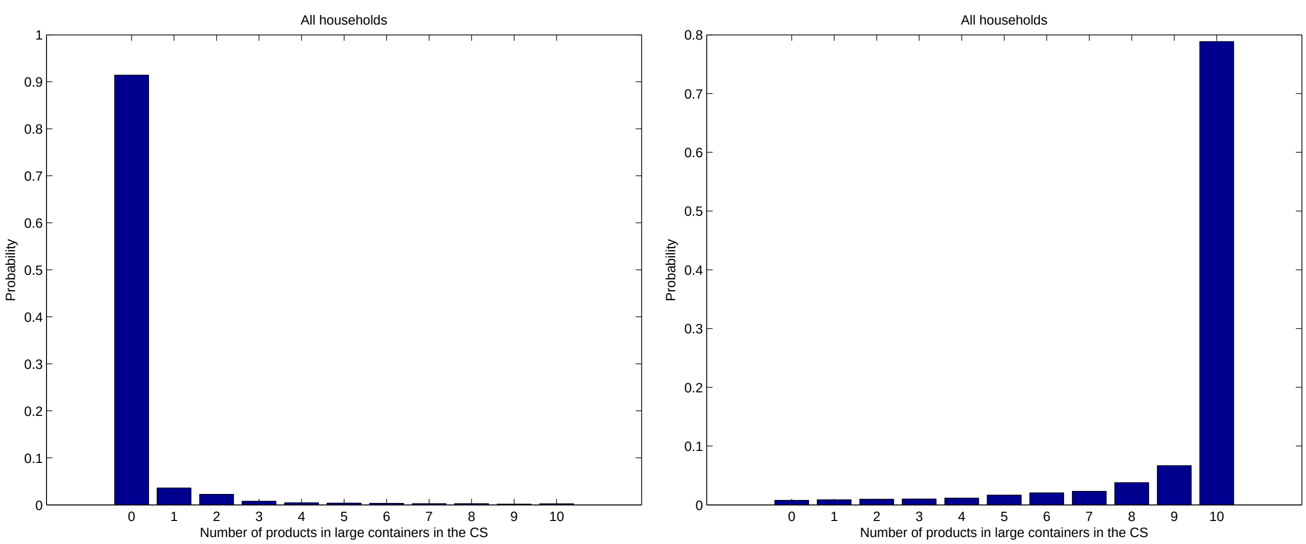
Figure 22: Average number of products considered



(a) Static Model (b) Dynamic Model

Note: Each panel reports the average number of products considered during a shopping trip in the designated model. Values are averaged over all shopping trips of the households in the specified income group. Panel (a) reports the averages using the estimates of the static model with a cash constraint reported in the third column of table 7; panel (b) reports the averages using the estimates of the dynamic model with a cash constraint reported in the second column of table 6.

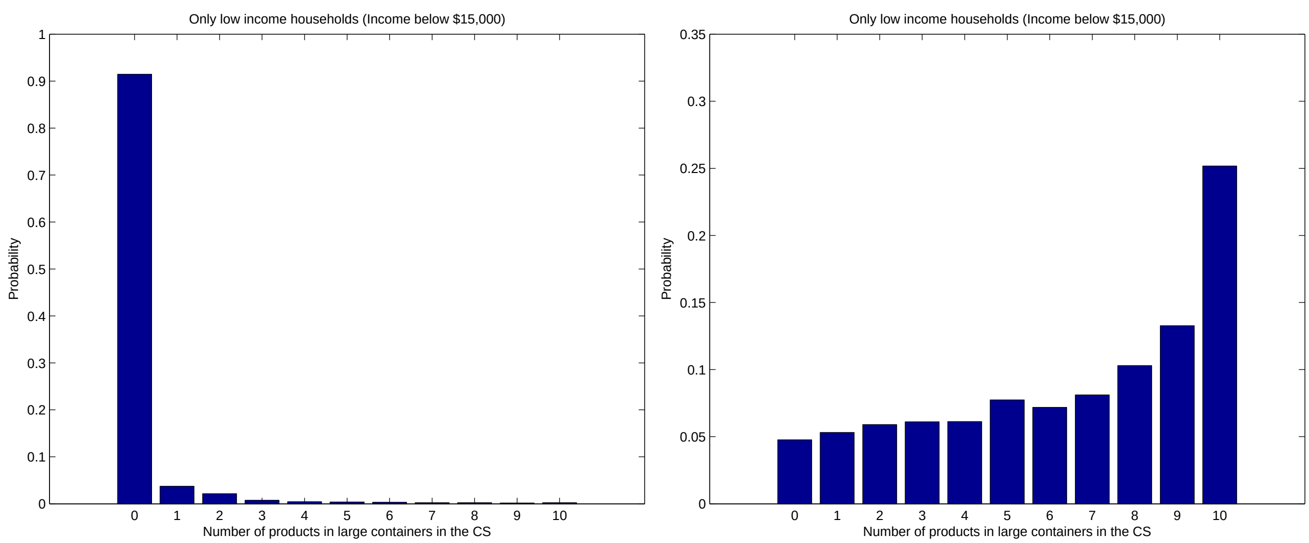
Figure 23: Distribution of number of products in large containers considered for all households (observation is a shopping trip)



(a) Static Model (b) Dynamic Model

Note: Each panel reports the probability distribution of the number of products in large containers considered during each shopping trip in the designated model. An observation is a shopping trip and the sample includes all households. Panel (a) reports the distribution using the estimates of the static model with a cash constraint reported in the third column of table 7; panel (b) reports the distribution using the estimates of the dynamic model with a cash constraint reported in the second column of table 6.

Figure 24: Distribution of number of products in large containers considered for low income households (observation is a shopping trip)



(a) Static Model (b) Dynamic Model

Note: Each panel reports the probability distribution of the number of products in large containers considered during each shopping trip in the designated model. An observation is a shopping trip and the sample includes only households with annual income lower than $15000. Panel (a) reports the distribution using the estimates of the static model with a cash constraint reported in the third column of table 7; panel (b) reports the distribution using the estimates of the dynamic model with a cash constraint reported in the second column of table 6.

11 Tables

Table 1: Summary Statistics

	mean	sd	skewness	min	max	p25	p50	p75
Shopping trip characteristics								
Time elapsed between trips (weeks)	0.673	0.655	1.193	0.000	7.000	0.000	1.000	1.000
Time elapsed btw purchases (weeks)	8.340	8.442	1.738	0.000	51.000	2.000	5.000	11.000
Volume bought in the last purchase (lb.)	10.083	7.481	4.498	0.032	160.000	6.250	8.000	12.500
Total Expenditure ($)	60.704	61.779	2.378	0.000	2279.212	17.735	40.674	83.308
Non-detergent expenditure($)	60.101	61.206	2.387	0.000	2279.212	17.538	40.289	82.501
Store Knowledge	0.956	0.206	-4.432	0.000	1.000	1.000	1.000	1.000
Mean Price ($/lb.)	1.064	0.268	4.733	0.360	9.630	0.903	0.987	1.116
Demographics								
Income (0000$)	4.511	2.722	0.477	0.500	10.000	2.250	4.000	6.196
Household's size	2.609	1.343	0.762	1.000	7.000	2.000	2.000	4.000
Purchase of liquid laundry detergent								
Price ($/lb.)	0.834	0.347	0.923	0.218	2.567	0.566	0.787	1.053
Size (lb.)	9.254	5.562	16.050	1.563	187.500	6.445	8.169	10.835
Number of brands purchased	3.836	3.019	1.155	1.000	18.000	1.000	3.000	6.000
Number of units purchased	32.271	53.405	7.426	1.000	1241.000	4.000	13.000	42.000
Household Brand HHI	0.605	0.306	0.143	0.101	1.000	0.337	0.535	1.000
Household Manufacturer HHI	0.662	0.285	-0.018	0.126	1.000	0.399	0.614	1.000
Shopping Behavior								
Total Expenditure ($)	68.857	39.177	1.647	8.926	441.815	40.731	61.229	87.609
Non-detergent expenditure ($)	68.076	38.802	1.645	7.165	436.511	40.277	60.549	86.928
Number of trips	245.336	296.434	1.944	1.000	2486.000	37.000	116.000	360.500
No.trips with purchases of laundry det.	21.891	30.567	3.355	1.000	359.000	3.000	10.000	30.000
Time elapsed btw purchases (weeks)	9.085	5.331	1.591	0.000	40.714	5.341	7.957	11.609
Time elapsed between trips (weeks)	0.902	0.402	1.728	0.000	5.000	0.623	0.850	1.093
Store visits								
Total number of stores visited	5.612	3.184	0.717	1.000	14.000	3.000	5.000	8.000
Store HHI (liq.laundry detergent)	0.712	0.266	-0.220	0.173	1.000	0.500	0.715	1.000
Brand Attributes								
Brand Price ($/lb.)	1.060	0.734	1.325	0.274	3.136	0.538	0.807	1.249
Brand Size (lb.)	7.470	3.808	0.690	0.725	17.990	4.417	6.968	9.484
Number of UPC's	25.697	38.729	1.840	1.000	156.000	2.000	5.000	27.000
Market Share	0.030	0.054	2.067	0.000	0.209	0.000	0.006	0.038
Brand HHI	0.123	–	–	0.123	0.123	0.123	0.123	0.123

Note: For Shopping Trip Characteristics an observation is a shopping trip (537,776 observations). For Brand Attributes an observation is a brand (33 observations). For the remaining statistics an observation is a household (2,192 observations). Store knowledge is the proportion of shopping trips that took place in a store visited by the household in the previous 12 weeks. Mean Price is the average price per pound of the brands available in the purchase occasion. Household Brand HHI is the sum of the square of the volume share of the brands bought by each household (i.e., concentration is evaluated within a given household). Household Manufacturer HHI is the sum of the square of the manufacturers' volume share by each household and Store HHI is the sum of the square of the expenditure share spent in each store by each household. Brand HHI is the sum of the square of the market share of each brand.

Table 2: Brand Market Shares and Promotional Activities

Brand	Size	Vendor	Share(Volume Sold)	Share (Revenues)	Price($/lb.)	Display	Feature
XTRA	Large	C&D	12.64%	6.68%	0.433	0.037	0.040
TIDE	Large	P&G	9.68%	14.19%	1.315	0.064	0.131
DYNAMO	Large	PHOENIX	8.75%	8.43%	0.777	0.017	0.043
XTRA	Medium	C&D	8.16%	4.37%	0.425	0.123	0.107
TIDE	Medium	P&G	7.55%	13.17%	1.439	0.085	0.108
PUREX	Large	DIAL	6.50%	5.26%	0.675	0.022	0.047
ARM	Large	C&D	5.85%	4.38%	0.730	0.036	0.046
PUREX	Medium	DIAL	5.66%	4.85%	0.733	0.089	0.087
Other	Medium		4.17%	4.37%	0.967	0.039	0.062
Other	Large		3.69%	3.31%	0.917	0.031	0.050
ALL	Medium	LEVER	3.65%	4.29%	0.982	0.075	0.091
ERA	Medium	P&G	3.47%	3.79%	0.947	0.116	0.098
ERA	Large	P&G	3.43%	3.56%	0.873	0.029	0.064
ARM	Medium	C&D	2.80%	2.59%	0.862	0.075	0.073
WISK	Medium	LEVER	2.52%	3.59%	1.170	0.069	0.089
ALL	Large	LEVER	2.47%	2.45%	0.883	0.033	0.026
PRIVATE	Large	PRIVATE	2.31%	1.30%	0.617	0.002	0.038
PRIVATE	Medium	PRIVATE	1.84%	1.10%	0.483	0.050	0.072
WISK	Large	LEVER	1.15%	1.57%	1.166	0.030	0.028
TIDE	Small	P&G	0.91%	2.26%	1.764	0.021	0.034
DYNAMO	Medium	PHOENIX	0.72%	0.79%	0.871	0.094	0.088
Other	Small		0.47%	0.86%	1.652	0.048	0.023
ALL	Small	LEVER	0.45%	1.02%	1.725	0.023	0.064
PUREX	Small	DIAL	0.40%	0.54%	1.101	0.080	0.090
ARM	Small	C&D	0.18%	0.28%	1.388	0.085	0.116
WISK	Small	LEVER	0.17%	0.42%	2.152	0.033	0.064
PRIVATE	Small	PRIVATE	0.16%	0.21%	1.295	0.026	0.030
ERA	Small	P&G	0.15%	0.25%	1.158	0.041	0.038
XTRA	Small	C&D	0.07%	0.06%	0.666	0.109	0.034
DYNAMO	Small	PHOENIX	0.05%	0.08%	1.305	0.122	0.087

Note: Column labeled Share (Qty Sold) are shares of volume sold in the sample, and column labeled Share (Revenues) are shares of revenues in the sample. The columns labeled Display and Feature present, respectively, the proportion of occasions a brand is displayed and featured. P&G = Procter and Gamble; C&D = Church and Dwight.

Table 3: Mean characteristics of laundry detergent bought by different households characteristics

		Package Size			Brand	Brand	Week of the month			
HH/Characteristics	Price/Lb.	Small	Medium	Large	Premium Brand	Private Label	1st	2nd	3rd	4th
Income:										
<20	0.817	0.143	0.528	0.329	0.388	0.045	0.244	0.276	0.257	0.223
20 - 45	0.798	0.093	0.586	0.321	0.401	0.042	0.232	0.275	0.265	0.228
40 - 65	0.767	0.055	0.602	0.343	0.405	0.042	0.228	0.273	0.272	0.227
>65	0.893	0.089	0.556	0.356	0.479	0.031	0.233	0.275	0.262	0.230
Family Size										
1	0.849	0.147	0.580	0.274	0.468	0.049	0.241	0.270	0.258	0.231
2	0.861	0.110	0.575	0.315	0.456	0.034	0.228	0.282	0.264	0.225
3	0.797	0.064	0.569	0.367	0.426	0.038	0.240	0.270	0.263	0.227
4	0.773	0.054	0.586	0.360	0.365	0.035	0.228	0.275	0.266	0.231
5	0.775	0.069	0.543	0.388	0.345	0.048	0.245	0.263	0.272	0.220
6	0.786	0.038	0.522	0.440	0.381	0.057	0.229	0.265	0.274	0.231
8	0.901	0.141	0.414	0.444	0.242	0.131	0.263	0.263	0.222	0.253
Age Cohort										
Born after 1955	0.800	0.068	0.569	0.362	0.385	0.041	0.234	0.272	0.266	0.227
Born before 1955	0.852	0.121	0.575	0.304	0.476	0.037	0.232	0.278	0.262	0.228

Table 4: Estimates of the effects on size-choice

	Laundry detergent				Shampoo			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	Size	Size	Size	Size	Size	Size	Size	Size
Income (000$)	0.00518***	0.0196***	-0.000935		-0.00162	-0.00338	-0.00352	
	(0.00101)	(0.00360)	(0.00141)		(0.00192)	(0.00583)	(0.00266)	
PeriodxBelowMedianIncome			-0.0164***				-0.00490	
			(0.00260)				(0.00468)	
Quarter Expenditure without good(000$)				0.0340***				0.00757
				(0.00724)				(0.0122)
PeriodxBelowQuarter Expenditure				-0.00813***				0.00743
				(0.00247)				(0.00499)
Household Size	0.0344***	0.111***	0.0344***	0.0401***	0.00735**	0.0184	0.00735**	-0.00289
	(0.00207)	(0.00734)	(0.00207)	(0.00732)	(0.00374)	(0.0112)	(0.00374)	(0.0159)
Last purchase quantity (lb.)	0.0278***	0.134***	0.0277***	0.00794***	0.100***	0.437***	0.100***	0.00449
	(0.000736)	(0.00289)	(0.000735)	(0.000443)	(0.00589)	(0.0245)	(0.00588)	(0.00499)
Time elapsed since the last purchase	-0.00174***	-0.00691***	-0.00167***	-0.000440	-0.00115**	-0.00403***	-0.00115**	-0.000178
	(0.000349)	(0.00126)	(0.000349)	(0.000362)	(0.000505)	(0.00150)	(0.000505)	(0.000591)
N	47986	47986	47986	47986	17016	17016	17016	17016
R^2	0.099		0.099	0.308	0.032		0.032	0.230

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 5: Estimates of the effects on size-choice

	Toothpaste			
	(1)	(2)	(3)	(4)
	Size	Size	Size	Size
Income (000$)	0.000473	0.00304	0.00155	
	(0.00167)	(0.00426)	(0.00230)	
PeriodxBelowMedianIncome			0.00289	
			(0.00413)	
Quarter Expenditure excluding good (000$)				0.0234**
				(0.0102)
PeriodxBelowQuarter Expenditure				-0.000792
				(0.00410)
Household Size	0.0439***	0.105***	0.0440***	0.0169
	(0.00334)	(0.00839)	(0.00334)	(0.0112)
Quantity bought during the last purchase (lb.)	0.440***	1.496***	0.440***	0.0173
	(0.0192)	(0.0555)	(0.0192)	(0.0135)
Time elapsed since the last purchase	-0.00131***	-0.00328***	-0.00132***	0.00287***
	(0.000468)	(0.00117)	(0.000468)	(0.000515)
N	31537	31537	31537	31537
R^2	0.044		0.044	0.255

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 6: Estimates for dynamic demand model with cash-constrained households

	Coeff.	SE	Coeff.	SE
Price Coefficient				
Constant	-0.6134	0.014	-0.6203	0.010
Family Size	0.0152	0.002	0.0143	0.003
Born before 1955	-0.0213	0.007	-0.0199	0.011
Income			0.0020	0.001
Display	0.6009	0.060	0.6025	0.060
Feature	0.3471	0.059	0.3479	0.059
Storage Cost				
Linear	0.0249	0.003	0.0250	0.002
Quadratic	-0.0081	0.001	-0.0082	0.001
Cash-Constraint				
Constant	9.8984	0.596	9.9067	0.640
Income	5.1025	0.636	5.1096	0.729
Dummy for 1st week of the month	-3.8873	0.660	-3.8836	0.744
Dummy for 2nd week of the month	5.6191	46.440	5.6192	52.471
Dummy for 3rd week of the month	-7.4571	0.567	-7.4518	0.587
Log-likelihood	19,636.70		19,635.50	
N	39277		39277	

Note: Specification includes brand and size fixed effects. Estimation is performed using a nested fixed point algorithm where the solution of the dynamic problem is nested within the parameter search. The rate of consumption of each household is the ratio of the total amount purchased to the overall time in the sample. The simulation of inventory starts with an initial guess and then updates inventory in each week using the observed purchases and the estimated consumption. Asymptotic standard errors are reported.

Table 7: Estimates for alternative models

	Alternative Models					
	Static w/o CC		Static w/o CC		Static w/ CC	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
Price Coefficient						
Constant	-0.6407	0.015	-0.6552	0.016	-0.1907	0.027
Family Size	0.0218	0.002	0.0201	0.002	0.0586	0.007
Born before 1955	-0.0181	0.007	-0.0150	0.007	-0.0517	0.016
Income			0.0041	0.001		
Display	0.6092	0.060	0.7107	0.077	0.5922	0.071
Feature	0.364	0.059	0.4620	0.075	0.1669	0.069
Display*Feature			-0.2218	0.114		
Cash-Constraint						
Constant					1.7174	0.099
Income					-0.0320	0.010
Dummy for 1st week of the month					0.0805	0.077
Dummy for 2nd week of the month					-0.0157	0.077
Dummy for 3rd week of the month					-0.058	0.078
Log-likelihood	19757.4		19749		19591.1	
N	39,277		39,277		39,277	

Note: Specification includes brand and size fixed effects. Asymptotic standard errors are reported.

Table 8: Counterfactual: All Households

	All Households	
	Cash-Constraints	No Cash-Constraints
Total Expenditure	4.1119	4.121
Average Expenditure By HH	6.8532	6.8683
Median Expenditure By HH	5.4665	5.4834
Percentile 25 Expenditure By HH	0	0
Percentile 75 Expenditure By HH	10.9417	10.9932
Total Volume Purchased	52.3750	52.4375
Average Volume Purchased By HH	8.72917	8.73958
Median Volume Purchased By HH	6.2500	6.250
Percentile 25 Volume Purchased By HH	0	0
Percentile 75 Volume Purchased By HH	12.500	12.500
Share Outside Option	0.9366	0.9365
Share Small Size	0.0082	0.00817
Share Medium Size	0.04475	0.0448
Share Large Size	0.0105	0.0105

Table 9: Counterfactual: Low-income households

Households with income below $5000

	Cash-Constraints	No Cash-Constraints
Total Expenditure	3.6809	4.2409
Average Expenditure By HH	6.1348	7.0681
Median Expenditure By HH	5.1012	5.7779
Percentile 25 Expenditure By HH	0	0
Percentile 75 Expenditure By HH	9.1072	11.0241
Total Volume Purchased	49.375	55.469
Average Volume Purchased By HH	8.22917	9.2448
Median Volume Purchased By HH	6.2500	6.2500
Percentile 25 Volume Purchased By HH	0	0
Percentile 75 Volume Purchased By HH	12.500	12.500
Share Outside Option	0.9409	0.9353
Share Small Size	0.00767	0.00758
Share Medium Size	0.04083	0.04400
Share Large Size	0.010583	0.01308

12 Appendix: Procedure to clean the raw data

The following describes the procedure we adopt to clean the raw data for liquid laundry detergent, but we follow the same procedure for the other product categories that we examine. Across all years 2001 to 2011, the store-level movement data for liquid laundry detergent tracks products, by Universal Product Code (UPC), 2,732 UPCs (10,557 for shampoo and 2,965 for toothpaste) in all, sold through 3147 stores. An observation is a store-week-UPC triple.

The household panel data are provided in two types of files. The first file, common to all product categories, contains information on all shopping trips made by each household in the panel, irrespective of what was purchased and which store was visited in that shopping trip. Here, an observation is a shopping trip (household-store-date triple). For each observation, the total expenditure is provided. Over the sample period, this file includes 6,805,177 shopping trips made by 14,809 panelists. The average duration of a household in the panel is 297 weeks (almost 7 years).

The second type of panel data file—a different file for each product category—details the complete purchase history by panelist, by UPC purchased within the product category, by shopping trip (store-date). An observation is then a household-store-date-UPC purchased. For liquid laundry detergent, over the sample period there are 267,177 purchases, 9,918 households and 1,320 UPCs, across 46 brands. For each observation, the number of purchased units and the average (imputed or actual) transaction price are provided.

The panel data is complemented with household income and other demographic characteristics. This information was updated in 2005, 2007, and 2011, and was current as of the date of the update.

Table 10 summarizes our data-cleaning procedure, applied to each product category separately, starting with the original number of observations (household-store-date-UPC purchased tuples) in the raw data. The procedure is as follows. We collapse household-store-date-UPC purchased level observations to the household-store-week level, taking the brand-size of the purchased UPC within the trip(s) that contains the largest number of consumption units and imputing as the transaction price of the unique brand-size the sales-weighted average price of the corresponding UPCs (i.e., sum of spending divided by sum of units purchased within the brand-size). We aggregate all shopping trips that occurred in the same store in the same week. We drop observations of households after a year with missing demographic information (income, household size or birth year of the head). For each household, we drop all observations in a year for which the household was not maintained in the panel for the entire 12 months. We drop households without shopping trips for at least one year. For each household, we drop the first 12 weeks of purchases (which we use to construct the household's "store knowledge"). We drop households whenever we cannot locate at least once the household-store purchase date in the household-store shopping trip file. We drop all observations before the first purchase and after the last purchase of the good (as we cannot compute the time between purchases). We drop households with at least an one-year lag between purchases. We drop households who did not purchase the good for at least one year. We thus consider, for each category, only household-store-date observations corresponding to periods in which the household is shopping regularly in the category. For laundry detergent, we drop observations after consumers switch from a liquid detergent to a non-liquid detergent. We further drop household-store-date observations when we cannot locate a consistent store-week-UPC in the store-level movement dataset.[16] We

[16]While the household panel data includes all shopping trips made by the household, some stores are excluded from the store-level movement dataset as they are not in the IRI system, thus we cannot recover store information. While the corresponding purchase observations are not used in estimation, they are used to construct the inventory level for each household in each week.

drop households without purchases in the cleaned sample.

The price of each brand-size available at a store is the sales-weighted average price of the corresponding UPCs. We assume that a brand-size pair is displayed (featured) if at least 50% of the sales-weighted UPCs that compose that brand-size are displayed (featured).

For some products listed for a store in a week, we observe zero purchases. This can happen if either the product was available but was not purchased, or the product was not available. For such a listed product with zero purchases, we impute a price equal to the maximum price we observe over the sample period for that product at that store, and we impute that the product was neither displayed nor featured. We thus assume that the product was available but was not purchased.

Table 10: Procedures to clean the raw data

	Purchase Data					
	Laundry detergent		Shampoo		Toothpaste	
Procedure	Obs dropped	Observations	Obs dropped	Observations	Obs dropped	Observations
Initial no. of obs. (hh-store-date-UPC purchased)		280,422		116,863		147,137
Collapse obs. to household-store-week level	29,572	250,850	20,858	96,005	19,038	128,099
No. purchases before merging with trip data		250,850		96,005		128,099

	Merged Shopping trip and Purchase Data					
	Laundry detergent		Shampoo		Toothpaste	
Procedure	Obs dropped	Observations	Obs dropped	Observations	Obs dropped	Observations
Initial no. of obs. (household-store-date)		6,805,177		6,805,177		6,805,177
Collapse observations to the household-store-week level	1,994,809	4,810,368	1,994,809	4,810,368	1,994,809	4,810,368
Drop obs. of hh after a year w/o demographic information	541,090	4,269,278	541,090	4,269,278	541,090	4,269,278
Drop observations from households who didn't make static	205,888	4,063,390	205,888	4,063,390	205,888	4,063,390
Drop households with a one-year lag between shopping trips	1,063,349	3,000,041	1,063,349	3,000,041	1,063,349	3,000,041
Drop each household's first 12 weeks in the dataset	107,033	2,893,008	107,033	2,893,008	107,033	2,893,008
Drop hh for whom we cannot locate purchase in trip file	492,475	2,400,533	270,440	2622568	168,805	2,724,203
Drop obs. before first purchase and after last purchase	403,680	1,996,853	1,038,453	1,584,115	730,249	1,993,954
Drop hh who did not purchase the good for at least 1 year	1,239,465	757,388	1,311,342	272,773	1,494,378	499,576
Drop obs after a hh first purchase of non-liquid detergent	162,769	594,619				
Drop obs. that cannot be located in the store-level data	53,361	541,258	28,613	244,160	47,295	452,281
Drop households without purchases in the cleaned sample	3,482	537,776	3,994	240,166	3,760	448,521
Final number of observations		537,776		240,166		448,521

13 Appendix: Additional Tables

Table 11: Summary Statistics - Shampoo

	mean	sd	skewness	min	max	p25	p50	p75
Shopping trip characteristics								
Time elapsed between trips (weeks)	0.668	0.656	1.211	0.000	7.000	0.000	1.000	1.000
Time elapsed btw purchases (weeks)	9.995	9.592	1.444	0.000	51.000	3.000	7.000	14.000
Volume bought in the last purchase (lb.)	1.388	1.121	6.205	0.062	26.185	0.844	0.938	1.625
Total Expenditure ($)	75.597	73.845	1.929	0.000	1984.066	21.479	51.899	107.122
Non-detergent expenditure($)	75.316	73.595	1.928	0.000	1979.703	21.360	51.705	106.770
Store Knowledge	0.957	0.203	-4.504	0.000	1.000	1.000	1.000	1.000
Mean Price ($/lb.)	4.392	0.705	0.274	1.009	8.908	3.891	4.355	4.839
Demographics								
Income (0000$)	4.645	2.716	0.392	0.500	10.000	2.250	4.000	6.278
Household's size	2.882	1.397	0.454	1.000	6.000	2.000	3.000	4.000
Purchase of liquid laundry detergent								
Price ($/lb.)	3.666	2.627	4.234	0.496	41.796	1.885	3.354	4.692
Size (lb.)	1.275	0.664	3.258	0.125	8.191	0.906	1.096	1.439
Number of brands purchased	4.050	4.036	1.891	1.000	24.000	1.000	2.000	5.000
Number of units purchased	19.361	39.862	7.995	1.000	797.000	2.000	5.000	19.000
Household Brand HHI	0.594	0.321	0.167	0.079	1.000	0.299	0.516	1.000
Household Manufacturer HHI	0.626	0.303	0.128	0.123	1.000	0.355	0.544	1.000
Shopping Behavior								
Total Expenditure ($)	78.478	42.747	1.485	7.361	422.237	47.585	69.905	99.653
Non-detergent expenditure ($)	78.143	42.610	1.485	6.656	421.241	47.475	69.477	99.480
Number of trips	191.215	270.947	2.657	1.000	2391.000	30.000	77.000	241.000
No.trips with purchases of laundry det.	13.548	22.299	2.905	1.000	162.000	1.500	4.000	14.000
Time elapsed btw purchases (weeks)	11.301	6.176	1.065	0.500	41.833	6.749	10.370	14.732
Time elapsed between trips (weeks)	0.885	0.377	1.243	0.000	3.000	0.617	0.837	1.079
Store visits								
Total number of stores visited	5.186	3.095	0.825	1.000	14.000	3.000	5.000	7.000
Store HHI (liq.laundry detergent)	0.782	0.255	-0.656	0.180	1.000	0.546	0.994	1.000
Brand Attributes								
Brand Price ($/lb.)	7.361	5.460	1.895	0.573	35.418	3.833	5.972	9.900
Brand Size (lb.)	1.127	0.725	5.372	0.167	7.125	0.750	1.052	1.328
Number of UPC's	15.111	31.680	3.330	1.000	175.000	1.000	3.000	9.500
Market Share	0.009	0.031	5.846	0.000	0.259	0.000	0.000	0.003
Brand HHI	0.110	–	–.	0.110	0.110	0.110	0.110	0.110

Note: For Shopping Trip Characteristics an observation is a shopping trip (240,166 observations). For Brand Attributes an observation is a brand (108 observations). For the remaining statistics an observation is a household (1,256 observations). Store knowledge is the proportion of shopping trips that took place in a store visited by the household in the previous 12 weeks. Mean Price is the average price per pound of the brands available in the purchase occasion. Household Brand HHI is the sum of the square of the volume share of the brands bought by each household (i.e., concentration is evaluated within a given household). Household Manufacturer HHI is the sum of the square of the manufacturers' volume share by each household and Store HHI is the sum of the square of the expenditure share spent in each store by each household. Brand HHI is the sum of the square of the market share of each brand.

Table 12: Brand Market Shares and Promotional Activities

Brand	Size	Vendor	Share(Vol. Sold)	Share(Rev.)	Price($/lb.)	Display	Feature
SUAVE	Medium	HELENE CURTIS	12.65%	6.58%	1.55	13.66%	9.57%
SUAVE	Large	HELENE CURTIS	12.58%	5.95%	1.22	12.87%	3.41%
Other	Large		9.53%	11.42%	3.83	5.43%	5.62%
Other	Medium		8.94%	13.03%	4.42	5.13%	9.15%
ALBERTO	Large	ALBERTO CULVER	7.09%	2.64%	0.91	0.00%	0.00%
PANTENE	Large	P&G	5.55%	8.34%	4.85	5.52%	6.67%
ALBERTO	Medium	ALBERTO CULVER	4.72%	1.94%	1.28	13.29%	7.69%
WHITE RAIN	Large	WHITE RAIN	4.65%	1.83%	1.12	16.20%	3.87%
TRESEMME	Large	ALBERTO CULVER	4.23%	2.74%	2.07	9.88%	7.82%
Other	Small		3.64%	8.97%	9.19	4.29%	6.70%
HEAD	Large	P&G	3.24%	4.86%	4.92	9.28%	1.91%
PANTENE	Medium	P&G	2.83%	5.00%	5.57	5.12%	7.09%
HEAD	Medium	P&G	2.68%	4.69%	5.47	4.57%	7.61%
LOREAL	Medium	LOREAL	2.48%	4.06%	5.02	5.12%	13.20%
WHITE RAIN	Medium	WHITE RAIN	2.42%	0.97%	1.29	14.29%	7.06%
LOREAL	Large	LOREAL	2.34%	3.06%			
CLAIROL	Large	P&G	2.30%	2.49%	4.32	3.93%	4.90%
CLAIROL	Small	P&G	1.92%	3.25%	5.19	7.46%	9.81%
CLAIROL	Medium	P&G	1.58%	2.02%	3.83	5.26%	11.82%
PRIVATE	Large	PRIVATE	1.19%	1.28%	2.29	5.92%	8.17%
PRIVATE	Medium	PRIVATE	0.98%	1.00%	3.37	5.81%	3.31%
PANTENE	Small	P&G	0.77%	1.32%	7.58	6.79%	9.93%
SUAVE	Small	HELENE CURTIS	0.65%	0.62%	2.89	3.34%	6.87%
LOREAL	Small	LOREAL	0.45%	1.00%	6.50	7.86%	11.21%
TRESEMME	Medium	ALBERTO CULVER	0.25%	0.31%	3.95	4.16%	8.86%
PRIVATE	Small	PRIVATE	0.24%	0.33%	5.13	5.02%	5.82%
HEAD	Small	P&G	0.09%	0.27%	9.22	0.66%	0.53%
TRESEMME	Small	ALBERTO CULVER	0.01%	0.03%	10.23	2.20%	1.28%

Note: Column labeled Share (Vol. Sold) are shares of volume sold in the sample, and column labeled Share(Rev.) are shares of revenues in the sample. The columns labeled Display and Feature present, respectively, the proportion of occasions a brand is displayed and featured. P&G = Procter and Gamble.

Table 13: Mean characteristics of shampoo bought by different households characteristics

		Package Size			Brand	Brand	Week of the month			
HH/Characteristics	Price/Lb.	Small	Medium	Large	Premium Brand	Private Label	1st	2nd	3rd	4th
Income:										
<20	3.212	0.149	0.538	0.313	0.370	0.041	0.222	0.269	0.270	0.239
20 - 45	3.545	0.156	0.563	0.281	0.425	0.025	0.240	0.274	0.258	0.228
40 - 65	3.427	0.153	0.584	0.263	0.416	0.027	0.228	0.271	0.266	0.234
>65	3.818	0.181	0.506	0.312	0.462	0.020	0.235	0.260	0.264	0.240
Family Size										
1	3.595	0.173	0.562	0.265	0.451	0.038	0.233	0.279	0.260	0.228
2	3.705	0.172	0.520	0.308	0.444	0.030	0.239	0.267	0.264	0.230
3	3.611	0.165	0.547	0.289	0.437	0.026	0.234	0.270	0.263	0.234
4	3.504	0.164	0.577	0.259	0.423	0.023	0.228	0.273	0.258	0.242
5	3.402	0.154	0.516	0.330	0.394	0.019	0.231	0.254	0.276	0.239
6	3.328	0.122	0.561	0.318	0.388	0.020	0.230	0.261	0.275	0.234
Age Cohort										
Born after 1955	3.491	0.160	0.555	0.285	0.420	0.026	0.229	0.270	0.265	0.236
Born before 1955	3.746	0.169	0.526	0.305	0.449	0.026	0.244	0.262	0.262	0.232

Table 14: Summary Statistics - Tootpaste

	mean	sd	skewness	min	max	p25	p50	p75
Shopping trip characteristics								
Time elapsed between trips (weeks)	0.640	0.639	1.198	0.000	7.000	0.000	1.000	1.000
Time elapsed btw purchases (weeks)	9.034	8.825	1.618	0.000	51.000	3.000	6.000	13.000
Volume bought in the last purchase (lb.)	0.530	0.397	6.422	0.031	8.125	0.363	0.400	0.663
Total Expenditure ($)	68.679	70.988	5.432	0.000	6022.183	18.757	45.932	96.296
Non-detergent expenditure($)	68.411	70.764	5.467	0.000	6022.183	18.661	45.750	95.927
Store Knowledge	0.957	0.204	-4.487	0.000	1.000	1.000	1.000	1.000
Mean Price ($/lb.)	10.111	1.100	0.300	3.644	17.600	9.321	10.109	10.833
Demographics								
Income (0000$)	4.635	2.698	0.428	0.500	10.000	2.250	4.000	6.426
Household's size	2.670	1.341	0.651	1.000	6.000	2.000	2.000	4.000
Purchase of liquid laundry detergent								
Price ($/lb.)	8.685	3.879	2.494	0.389	52.408	6.564	7.986	9.842
Size (lb.)	0.478	0.211	3.595	0.047	3.375	0.365	0.431	0.543
Number of brands purchased	2.744	1.899	1.252	1.000	11.000	1.000	2.000	4.000
Number of units purchased	26.549	41.113	3.190	1.000	391.000	3.000	9.000	34.000
Household Brand HHI	0.687	0.271	-0.046	0.152	1.000	0.456	0.634	1.000
Household Manufacturer HHI	0.703	0.260	-0.049	0.210	1.000	0.481	0.662	1.000
Shopping Behavior								
Total Expenditure ($)	74.142	42.197	1.468	5.766	441.815	44.133	65.409	95.508
Non-detergent expenditure ($)	73.834	42.068	1.469	5.370	441.369	43.914	65.269	95.168
Number of trips	275.843	355.736	1.999	1.000	2436.000	38.000	118.000	383.000
No.trips with purchases of laundry det.	19.395	28.136	2.943	1.000	327.000	2.000	7.000	26.000
Time elapsed btw purchases (weeks)	10.939	5.790	1.555	0.000	49.500	6.875	9.822	13.536
Time elapsed between trips (weeks)	0.878	0.381	1.311	0.000	3.200	0.616	0.819	1.065
Store visits								
Total number of stores visited	5.705	3.368	0.766	1.000	14.000	3.000	5.000	8.000
Store HHI (liq.laundry detergent)	0.752	0.265	-0.494	0.143	1.000	0.511	0.836	1.000
Brand Attributes								
Brand Price ($/lb.)	15.513	11.853	1.338	2.231	54.968	7.399	12.088	20.999
Brand Size (lb.)	0.425	0.443	5.416	0.035	3.250	0.250	0.365	0.493
Number of UPC's	23.163	68.001	4.504	1.000	395.000	1.000	3.000	16.000
Market Share	0.020	0.076	4.540	0.000	0.428	0.000	0.000	0.002
Brand HHI	0.297	–	–	0.297	0.297	0.297	0.297	0.297

Note: For Shopping Trip Characteristics an observation is a shopping trip (448,521 observations). For Brand Attributes an observation is a brand (49 observations). For the remaining statistics an observation is a household (1,626 observations). Store knowledge is the proportion of shopping trips that took place in a store visited by the household in the previous 12 weeks. Mean Price is the average price per pound of the brands available in the purchase occasion. Household Brand HHI is the sum of the square of the volume share of the brands bought by each household (i.e., concentration is evaluated within a given household). Household Manufacturer HHI is the sum of the square of the manufacturers' volume share by each household and Store HHI is the sum of the square of the expenditure share spent in each store by each household. Brand HHI is the sum of the square of the market share of each brand.

Table 15: Brand Market Shares and Promotional Activities

Brand	Size	Vendor	Share(Vol.Sold)	Share(Rev.)	Price ($/lb.)	Display	Feature
COLGATE	Large	COLGATE	27.14%	23.40%	6.71	7.27%	13.52%
CREST	Large	P&G	18.32%	16.53%	6.91	3.98%	10.14%
COLGATE	Medium	COLGATE	11.68%	11.26%	7.57	8.27%	16.77%
CREST	Medium	P&G	10.96%	10.72%	7.81	4.30%	14.26%
AQUAFRESH	Large	GSK	5.12%	3.81%	4.65	5.54%	4.10%
COLGATE	Small	COLGATE	4.02%	5.44%	10.75	4.75%	10.63%
AQUAFRESH	Medium	GSK	3.25%	3.00%	7.41	2.80%	7.81%
CREST	Small	P&G	2.49%	3.42%	10.90	1.93%	4.80%
MENTADENT	Large	C&D	1.98%	2.56%	9.54	0.62%	2.09%
SENSODYNE	Small	GSK	1.82%	4.81%	20.93	1.06%	7.02%
Other	Large		1.67%	1.56%	3.53	60.47%	2.10%
SENSODYNE	Large	GSK	1.34%	2.63%	13.93	48.61%	0.76%
AQUAFRESH	Small	GSK	1.27%	1.54%	9.87	1.56%	3.15%
Other	Medium		1.22%	1.11%	7.46	3.89%	8.02%
PEPSODENT	Medium	C&D	1.07%	0.44%	3.53	4.23%	8.87%
PEPSODENT	Large	C&D	0.99%	0.39%	2.91	15.33%	5.44%
AIM	Large	C&D	0.89%	0.33%	2.73	3.94%	4.71%
Other	Small		0.89%	2.27%	20.87	1.62%	2.67%
AIM	Medium	C&D	0.87%	0.37%	3.41	3.52%	7.98%
ARM	Large	C&D	0.69%	0.92%	6.15	0.00%	0.00%
ARM	Small	C&D	0.66%	1.00%	13.66	0.76%	3.11%
MENTADENT	Medium	C&D	0.60%	1.01%	11.75	1.63%	5.35%
ARM	Medium	C&D	0.52%	0.64%	9.62	0.89%	5.73%
MENTADENT	Small	C&D	0.19%	0.44%	17.53	1.66%	6.01%
PRIVATE	Large	PRIVATE	0.15%	0.13%	3.93	14.50%	3.45%
PRIVATE	Medium	PRIVATE	0.07%	0.04%	4.72	9.07%	4.41%
SENSODYNE	Medium	GSK	0.07%	0.14%	16.23	2.06%	3.11%
PRIVATE	Small	PRIVATE	0.06%	0.10%	11.05	3.25%	5.73%
AIM	Small	C&D	0.00%	0.00%	8.28	0.00%	0.20%

Note: Column labeled Share (Vol. Qty) are shares of volume sold in the sample, and column labeled Share(Rev.) are shares of revenues in the sample. The columns labeled Display and Feature present, respectively, the proportion of occasions a brand is displayed and featured. P&G = Procter and Gamble; GSK = GLAXOSMITHKLINE; C&D = Church and Dwight.

Table 16: Mean characteristics of toothpaste bought by different households characteristics

HH/Characteristics	Price/Lb.	Package Size Small	Package Size Medium	Package Size Large	Brand Premium Brand	Brand Private Label	Week of the month 1st	Week of the month 2nd	Week of the month 3rd	Week of the month 4th
Income:										
<20	8.073	0.260	0.382	0.357	0.302	0.003	0.254	0.282	0.246	0.218
20 - 45	8.555	0.217	0.418	0.365	0.381	0.003	0.245	0.282	0.249	0.224
40 - 65	8.841	0.231	0.411	0.358	0.404	0.003	0.256	0.274	0.248	0.223
>65	8.977	0.230	0.367	0.403	0.371	0.003	0.245	0.279	0.257	0.218
Family Size										
1	8.662	0.292	0.408	0.299	0.292	0.001	0.252	0.274	0.248	0.226
2	8.864	0.246	0.386	0.368	0.406	0.004	0.250	0.286	0.247	0.217
3	8.722	0.218	0.390	0.392	0.389	0.003	0.258	0.273	0.252	0.218
4	8.354	0.207	0.417	0.376	0.346	0.004	0.243	0.281	0.251	0.225
5	8.646	0.171	0.397	0.432	0.375	0.001	0.243	0.264	0.262	0.231
6	8.319	0.144	0.393	0.463	0.354	0.003	0.224	0.281	0.266	0.229
Age Cohort										
Born after 1955	8.563	0.205	0.402	0.392	0.374	0.003	0.248	0.279	0.250	0.223
Born before 1955	8.809	0.256	0.391	0.353	0.372	0.003	0.249	0.280	0.252	0.219