Less is More Expensive: Bulk Buying and Cognitive Costs[*†]

Christoph Bauner[‡] Mallick Hossain[§]

April 24, 2025

Abstract

Increasing the salience of unit prices can reduce consumption inequality. Using NielsenIQ data, we show that low-income households forgo savings by not buying in bulk. We estimate that low-income households could save 5% on groceries if they bought in bulk like high-income households. Using novel data on state-level unit-price regulations, we find that cognitive costs discourage households from bulk buying. Mandating unit price display, a policy adopted by nine states, may reduce cognitive costs and increase the salience of unit prices. This policy may help close the "bulk buying gap" by 36% because low-income households react most strongly to it.

1 Introduction

Poor households often pay more for goods and services, ranging from tens or hundreds of dollars in fees for banking services to thousands of dollars in higher interest for mortgages and car loans (Fellowes, 2006). Often, these higher costs are attributed to discrimination or search costs. In this paper, we show that even when discrimination is nearly impossible and search costs are vanishingly small, low-income households pay more for goods. Specifically, we find that low-income households pay more for everyday grocery items because they do not buy in bulk as frequently as other households, even when the same product is available for a lower unit price in the store.[1] This pattern is especially common for storable, essential goods (e.g., toilet paper or

[*]The views expressed here are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This article is based on a chapter of Mallick Hossain's dissertation. He thanks his advisers, Katja Seim, Holger Sieg, Aviv Nevo, and Sarah Moshary for their advice and constructive criticism. Thanks to Mike Abito, Francesco Agostinelli, Emek Basker, Minsu Chang, Chris Cronin, Frank DiTraglia, Eileen Divringi, Hanming Fang, Jim Ferry, Ishan Ghosh, Joao Granja, Kathleen Hui, Ben Hyman, Jeff Lin, Igor Livshits, Paolo Martellini, Davin Reed, Claudia Sahm, Paul Sangrey, Andrew Shephard, Petra Todd, Anna Tranfaglia, Sofia Villas-Boas and Keith Wardrip for their comments. We also thank Sofia Villas-Boas and audiences at the University of Massachusetts Amherst, Purdue University, the International Industrial Organization Conference 2024, and the Banco de la República (Colombia) Workshop on Industrial Organization and Development 2024.

[†]Researchers' own analyses derived based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researchers and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

[‡]University of Massachusetts Amherst; cbauner@umass.edu

[§]Federal Reserve Bank of Philadelphia; mallick.hossain@phil.frb.org

[1]"Bulk" refers to package sizes in the top two quintiles of all products available within a particular product category (Griffith et al., 2009). This definition is further discussed in Section 3.

detergent) for which demand is relatively fixed and predictable. We show that policies mandating the display of unit prices, which increase the salience of unit prices and reduce cognitive costs, encourage households to use bulk discounts.[2]

Grocery purchases account for a sizable share of a household's discretionary spending, especially for the lowest-income households (BLS, 2019). To save money, households can wait for sales, redeem coupons, purchase generic brands, or even increase home production (Griffith et al., 2009; Aguiar and Hurst, 2005). Bulk buying is another way for households to save money. Bulk discounts are almost always available without the need to wait for a sale, possess a coupon, or choose a less-preferred brand. Even though households buy more on a given shopping trip, they pay lower unit prices and reduce their overall spending.

We show that, despite the substantial savings available from quantity discounts, low-income households are less likely to buy in bulk than high-income households.[3] This gap is particularly large for non-food items where about 40% of low-income purchases are bulk compared to 50% for high-income households. Kunreuther (1973) provides the first evidence of this "bulk buying gap" for a few specific products, and Orhun and Palazzolo (2019) expand this finding to a whole product category. Since households purchase a variety of products when shopping, we show that the bulk buying gap exists across the full range of product categories that households purchase.

We find that cognitive costs contribute to this bulk buying gap. First, the cognitive costs of assessing price differences across products can prevent households from making economical decisions (Mitchell et al., 2003). Providing price information reduces the effort needed to compare prices, and households change their purchase decisions when such information is more salient (Chetty et al., 2009; Bogomolova and Jarratt, 2016). To be clear, we make no claim about the cognitive differences across households. Rather, posting unit-prices reduces the cognitive burden on all households. Because low-income households are more price-sensitive than their wealthier counterparts, they react more strongly to such a lowering of the cognitive cost and, therefore, the bulk-buying gap narrows.

Posting unit prices reduces the cognitive costs of comparing unit prices across different products. Only nine states require the display of unit prices, and no study has evaluated the impact of these regulations on consumer behavior. We provide the first nationwide study of the impact of displaying unit prices on bulk purchasing and find that households are significantly more likely to buy in bulk when retailers are mandated to display unit prices.

For our analysis, we combine household- and store-level datasets to study income heterogeneity in bulk buying. NielsenIQ's Consumer Panel data are a nationally representative panel survey of household grocery purchases, and NielsenIQ's Retail Scanner data are a national panel of weekly UPC-level sales data from over 30,000 stores. We construct a new dataset of state-level per-unit pricing regulations, including a measure of

[2]These policies impose strict standards on how unit prices are calculated and displayed. Without such policies, retailers may omit unit pricing, display it in microscopic font, or even calculate the unit price based on different units within the same category (say per ounce for one bottle of detergent and per pint for a different bottle).

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

regulatory stringency. As a result, we have a comprehensive view of a household's possible product choices, available price information, and resulting expenditures.

We find that low-income households could realize substantial savings from buying in bulk at the same rate as high-income households. This is based on estimating the average bulk discount for each product category based on NielsenIQ's weekly store-level price and product data. The average discount across all product categories is such that a package that is twice as large will have a 30% lower unit price. Then, we estimate how much each household buys in bulk using NielsenIQ's household-level purchase data. Given each product-category-specific bulk discount and household-level bulk buying, we predict how much low-income households could save if they increased their bulk buying intensity to match that of high-income households. We find that low-income households would reduce their annual grocery expenditures by 5% if they bought in bulk like high-income households, saving an aggregate of $5.4 billion annually.

We then employ a difference-in-differences model to determine how much cognitive costs affect bulk buying. Our analysis exploits the fact that a significant share of households in the NielsenIQ Consumer Panel switch between regulatory regimes when they move from one state to another. This design allows us to include household fixed effects in the regression. Movers purchase more than 1 percentage point more in bulk when in a state with unit price regulation than the same household would in states without such a rule.

Finally, we construct a discrete-choice model of consumer purchasing behavior to quantify consumer preferences and disentangle the contribution of cognitive costs to the bulk buying decision. We estimate this model using data on toilet paper purchases. Households choose a product based on price, quantity, quality, and package size, which serves as a proxy for storage costs. We can separate preferences for quantity from size preferences because we demonstrate that toilet paper comes in varying "concentrations."[4] We allow state-level unit pricing mandates to affect a household's unit price sensitivity. From this demand model, we simulate household responses to the counterfactual of universally posting unit prices.

Our model predicts that requiring stores to post unit prices would reduce the bulk buying gap in package size purchased between high- and low-income households by 36%. As a result of this policy, households would buy larger quantities of toilet paper and pay lower unit prices. Universally displaying unit prices would encourage households to better utilize quantity discounts by reducing cognitive costs, increasing bulk buying, and helping consumers save money.

The rest of the paper is structured as follows. Section 2 describes the data. Section 3 documents new facts of quantity discounting. Section 4 presents evidence of cognitive costs contributing to the bulk buying gap. Section 5 introduces the model. Section 6 presents estimation and counterfactual results. Section 7 concludes.

[4]Toilet paper "concentration" roughly amounts to differences in the number of sheets per roll. As a concrete example, a 12-pack of Charmin can contain anywhere from about 140 sheets per roll to about 430 sheets per roll. We discuss this at length in Section 5.

2 Data

In this section, we describe the datasets used for our analysis and give a brief overview of their respective features. NielsenIQ's Consumer Panel data provide information on households' shopping and purchasing decisions. NielsenIQ's Retail Scanner data provide information on weekly product assortments and prices. The new regulatory dataset we construct contains information on state-level regulations regarding the display of per-unit pricing. By combining these data, we have a comprehensive view into a household's possible product choices, available price information, and their resulting purchase decisions.

2.1 NielsenIQ Consumer Panel Data

We use the NielsenIQ Consumer Panel dataset from 2004–2019. This dataset is a panel of about 195,000 unique households. We observe about 40,000 households each year from 2004–2006 and about 60,000 households each year from 2007–2019. Households scan all items that they purchase and then input information about quantities, prices, date of purchase, and store. NielsenIQ retains about 80% of its panel from year to year with the mean and median tenure of a household being 4.7 and 3 years, respectively.

We consider food, drink, and non-food grocery (e.g., paper towels, toilet paper, detergent) purchases made at grocery stores, discount stores, dollar stores, warehouse clubs, and drugstores. These outlets account for over 90% of household expenditures in these categories. We exclude alcohol, tobacco, health, and general merchandise products from our analysis since these products (e.g., cigarettes, painkillers) may have different consumption patterns than grocery products or are not suited for bulk purchases (e.g., printers, cookware, linens). We also exclude households with a student or military head of household as well as those with an annual income of less than $5,000 and those living in mobile homes. Only about 7% of households are excluded, and we use the remaining 181,000 households for our analysis. See Appendix A.1 for further details of sample construction.

NielsenIQ computes projection weights to ensure their sample is nationally representative. Weights are calculated to match population moments based on household size, income, age, race, ethnicity, education, occupation, and presence of children. All analyses use these projection weights unless otherwise stated. NielsenIQ groups household income into 16 different income bins. Due to the large number of bins, in tables and parts of the text, we will report differences by income quartiles. However, where possible (especially in graphs), we will report estimates for each income bin. Table 1 presents descriptive statistics for households in the sample.

2.2 NielsenIQ Scanner Data

The NielsenIQ Scanner data contain average weekly prices and volume sold of individual products at about 49,000 stores from about 146 retail chains between 2006–2019. Average prices are weighted by the volume sold. Only products with positive sales in a given week are recorded. We match the Retail Scanner data with

Table 1: Summary Statistics of NielsenIQ Households

Variable	Mean	SD	25th Pctile	75th Pctile
Household income ($000s)	58.03	31.37	27.5	85
Household size	2.56	1.45	1	3
Age	52.70	14.44	41.5	63
College educated	0.39	0.49	0	1
Child present	0.32	0.47	0	1
Married	0.51	0.50	0	1
N (Household-Years)	849,145			
N (Households)	181,481			

the Consumer Panel data based on store identification numbers and purchase dates. By matching the two datasets, we recover the set of products available to a household and the product it chose to purchase.

2.3 Unit Pricing Regulations

We compile a novel dataset on state-level regulations regarding the display of unit prices. The data are based on annual regulatory updates aggregated in Handbook 130 published by the National Institute of Standards and Technology (NIST, 2019). We cross-check this information with state regulatory codes and state officials to ensure accuracy. This data include information on which states have regulations, when they were adopted, and how stringent these regulations are. More details are discussed in Section 4.

3 Stylized Facts

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

3.1 Quantity Discount Prevalence

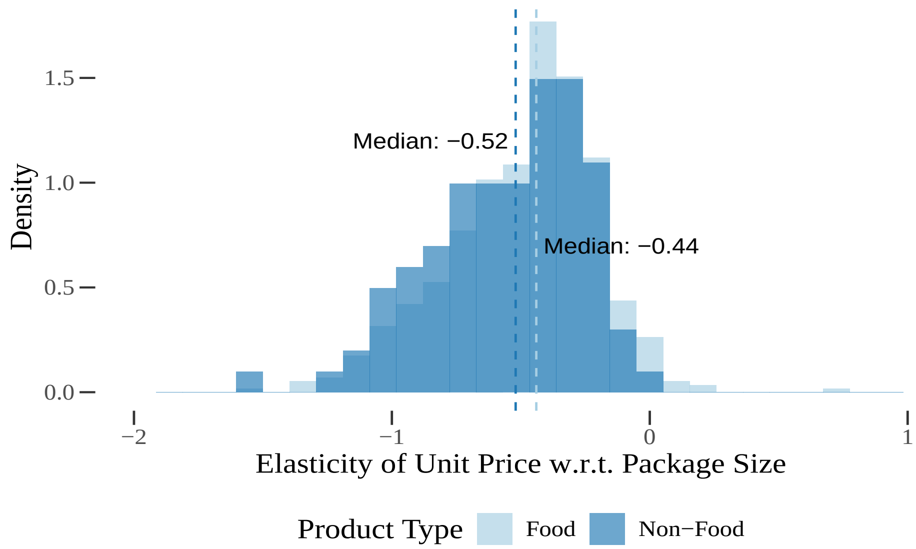
Quantity discounts are a specific form of non-linear pricing in which unit prices decrease as package size increases. To establish the prevalence and magnitude of quantity discounts, we use NielsenIQ's Retail Scanner data from 2019. We estimate quantity discounts separately for 655 product categories using the following regression:

$$\ln(P)_{ibm} = \alpha + \beta \ln(Size)_{ibm} + \lambda_{bm} + \epsilon_{ibm}, \tag{1}$$

where P_{ibm} is the unit price (package price divided by package size) of product i from brand b purchased in market m (defined as a store-week). $Size_{ibm}$ is the item's package size, which is the number of units included in a UPC (e.g., quart, square feet, count, pound). λ_{bm} is a brand-store-week fixed effect. Variation in unit prices across package sizes within the same brand-store-week identify β.[5] If retailers offer quantity discounts, then β will be negative.

Figure 1 plots the distribution of β across product categories (statistically insignificant betas are set to zero); 87% of all product categories have a statistically significant and negative β, and non-food items generally have larger discounts than food items.[6] The median β is -0.52 for non-food products, which means that a package that is double the size will have a 30% lower unit price.[7] This discount is larger than the median β for food items (-0.44).[8] The size and near-universality of quantity discounts suggest they offer substantial savings to households without sacrificing consumption.[9]

Figure 1: Distribution of Bulk Discounts by Product Type



Note: Figure plots the distribution of coefficients from a regression of log unit price on log package size (Equation (1)) for individual product categories. Regression controls for store-brand-week fixed effects. **Source:** NielsenIQ Retail Scanner (2019).

3.2 Bulk Purchasing

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

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

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

[7] Doubling the package size requires converting the log-point approximation to actual percent changes: exp(-0.52 * ln(2)) = 0.697.

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

[9] For a comparison of quantity discounts with coupons, see Appendix A.2.

bulk buying will likely differ between food and non-food items. Because of these differences, we analyze food and non-food products separately. Following the literature, we classify a product as "bulk" if it is in the top two quintiles of the size distribution for that product category (Griffith et al., 2009). This definition can easily be applied across all product categories without needing to motivate specific cutoffs for each unit of measure and product category. Furthermore, the top two quintiles of the size distribution capture large sizes that are available at grocery stores and mass merchandisers.[10] Then, for each household, we compute the expenditure share of bulk purchases of food and non-food items. We then regress this "bulk share" on household income and other household characteristics that could affect consumption patterns and may be correlated with income and plot the income coefficients. The equation below is estimated on food and non-food purchases separately:

$$BulkShare_{imt} = \alpha + \sum_{q} \beta^q Income_{imtq} + \gamma X_{imt} + \lambda_m + \lambda_t + \epsilon_{imt}, \tag{2}$$

where $BulkShare_{imt}$ is household i's share of bulk purchases in market m in year t (a market is a county). $Income_{imtq}$ is a dummy equal to 1 if household i's income in year t falls into income bin q. X_{imt} consists of household characteristics (age, household composition, marital status, education, housing type, tract-level vehicle access).[11] Year- and market-fixed effects are captured by λ_t and λ_m.

Figure 2 illustrates that bulk purchases comprise a 10 percentage point larger share of non-food expenditures for households making over \$100,000 compared to those making \$5,000–\$8,000. As income increases, bulk purchases make up an increasing share of expenditures. For food items, there is a more muted increase of one percentage point across income groups.

The 10 percentage point gap is quite large. For average households making between \$5,000 and \$8,000, 39.6% of their non-food grocery spending is on bulk packages. Hence, households making over \$100,000 are 26% more likely to buy in bulk relative to the lowest-income group.

These patterns are consistent with high-income households buying in bulk, obtaining low unit prices, and consuming out of storage. Given the existence of quantity discounts, larger packages generally correspond to lower unit prices. The fact that low-income households are less likely to buy these storable items in bulk suggests that some obstacles may prevent them from buying and storing large packages.[12]

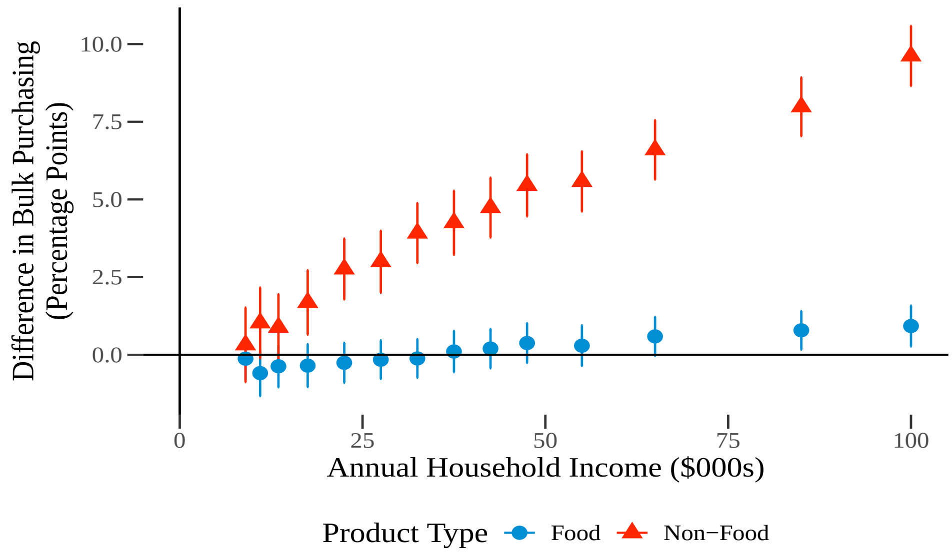
Because the bulk buying gap is largest for non-food products, the rest of this paper focuses on non-food products. These products are ideal for analyzing bulk purchasing because they isolate the key features that make bulk buying and quantity discounts attractive for households. Primarily, households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

[10] Using only the top quintile risks capturing only the largest sizes that may only be available at warehouse clubs.

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

[12] This relationship persists across most categories. Appendix A.3 shows the same pattern for a few popular categories.

Figure 2: Bulk Purchasing by Household Income and Product Type



Note: Figure plots the income bin coefficients from Equation (2), which regresses the share of annual purchases that were bulk packages on household characteristics as well as market and year fixed effects. NielsenIQ projection weights are used to ensure national representativeness. Households making $5k–$8k are the reference group. Vertical lines indicate 95% confidence intervals.
Source: NielsenIQ Consumer Panel (2004–2019)

many food products (Aguiar and Hurst, 2005, 2007).

3.3 Savings from Bulk Buying

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

$$\ln(AvgSize)_{imt} = \alpha + \sum_q \beta^q Income_{imtq} + \gamma X_{imt} + \lambda_m + \lambda_t + \epsilon_{imt}, \tag{3}$$

where $AvgSize_{imt}$ is the quantity-weighted average package size purchased by household i in market m in year t, where a market is a county.[13] $Income_{imtq}$ is an indicator for a household's income quartile. X controls for household characteristics. Market and year fixed effects are included through λ_m and λ_t.

In this regression, β^q gives the average log-difference between the package size purchased by a household in income quartile q and the lowest-income quartile (households making less than $25,000).[14] To compute savings, we multiply this average difference in package size purchased by the category-specific quantity discount estimated in Section 3.1. For example, high-income households buy 26% larger packages of toilet paper, which has a quantity discount of 0.253. Therefore, low-income households could save $0.26 \times 0.253 = 0.066$ or

[13]Average package size is weighted by quantity to account for the fact that an unweighted average would favor small packages.
[14]We use quartiles to reduce the number of income bins from 15 to 4, but results hold at more granular levels. Disaggregated results are available upon request.

6.6% from buying big packages like high-income households do. Aggregating across all categories in which high-income households buy larger packages gives an estimated savings of 5%, or $207, per year.[15,16]

Saving 5% on these common household purchases is substantial for low-income households. For the bottom quintile of the income distribution, these items account for 30% of their discretionary spending compared to 19% for the top quintile of the distribution.[17] If the 24.4 million households making under $25,000 were to obtain these savings, that would be an overall savings of $5.4 billion annually, assuming no supply-side response.[18] For context, this is equal to 8% of the $68 billion federal Supplemental Nutrition Assistance Program budget in 2017 (USDA, 2019). These potential savings do not require low-income households to buy more over the course of the year because buying in bulk does not necessarily change how much households *consume*. It just changes how much they *buy* at one time.

4 Cognitive Costs and Bulk Buying

In this section, we show that cognitive costs affect the bulk buying decision. To do this, we use plausibly exogenous variation to estimate the causal impact of unit pricing regulation on bulk purchasing. Since the biggest differences in bulk buying are for non-food grocery items, all analysis is restricted to non-food products.

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

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

Using annual regulatory updates published by the National Institute of Standards and Technology (NIST), we compile state-level regulations on unit pricing (NIST, 2019). For states with regulations, we cross-check

[15]This averages only across categories in which high-income households buy larger packages. There are some categories, such as septic tank cleaners, in which high-income households buy in smaller packages. Imposing that low-income households buy the same average size across *all* categories reduces projected savings to about 2%.

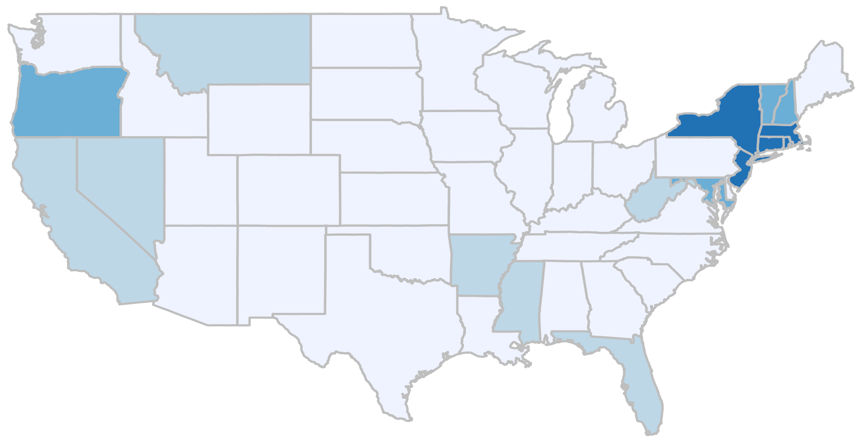
[16]The first-best calculations of savings would identify the product with the lowest unit price given a household's brand and store choice and compute savings based on that product. This estimate will likely be substantially higher than what we computed, so we view the estimated 5% savings as a conservative estimate of potential savings. See Appendix A.4 for calculations of savings on popular product categories.

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

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

NIST's designation with state regulatory codes and consult with state officials to ensure accuracy. Figure 3 shows that, as of 2019, 16 states have regulations on the display of unit prices and 34 have no regulations.[19]

Figure 3: Unit Price Regulations by State (2019)



Note: Figure plots whether or not a state has regulations in place governing the display of unit prices as of August 1, 2019. "No Reg" denotes that no regulations are in effect. "Vol. Disp" denotes states where regulations apply if retailers choose to display unit prices. "Mand. Disp" denotes states where all retailers must display unit prices. "Mand. Disp, Strict" denotes states where strict display formatting requirements are in effect. **Source:** NIST Handbook 130

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

$$BulkShare_{it} = \alpha + \beta_1 Reg_{it} + \gamma X_{it} + \lambda_t + \epsilon_{it}, \tag{4}$$

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

Since 2004, no state has modified its regulations on unit prices, so the coefficient on unit pricing regulation is identified from cross-sectional variation between states that have regulations and those that do not.[20,21] Column (1) of Table 2 reveal that bulk purchasing is 3.4 percentage points higher in states with unit price regulations compared to states without unit price regulation, even after controlling for household

[19]Summary statistics of these groups are reported in Appendix Table A6.

[20]Because in out data there is no time variation in regulations, we cannot include state fixed effects in the estimation.

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

characteristics and year fixed effects. Furthermore, Column (2) shows that there is a distinct pattern by income in that unit price laws are associated with higher bulk buying overall, but the effect is even stronger for higher-income households. Quantitatively, the lowest income quartile has 2.5 percentage points higher bulk buying in states with unit price posting laws, but the highest-income quartile has 4 percentage points higher bulk buying.

Table 2: Correlation of Bulk Buying and Demographics (All Products))

Model:	(1)	(2)	(3)	(4)
Variables				
Regulation	0.0338***	0.0251***		
	(0.0005)	(0.0014)		
Regulation × \$25-\$50k		0.0037**		
		(0.0017)		
Regulation × \$50k-\$100k		0.0121***		
		(0.0016)		
Regulation × >\$100k		0.0157***		
		(0.0018)		
Voluntary			0.0456***	0.0059***
			(0.0006)	(0.0008)
Mandatory			0.0365***	0.0364***
			(0.0013)	(0.0013)
Mandatory Strict			0.0270***	0.0269***
			(0.0008)	(0.0008)
Demographics	Y	Y	Y	Y
Omit CA	N	N	N	Y
Fit statistics				
Observations	846,543	846,543	846,543	773,328
Adjusted R^2	0.05805	0.05818	0.04481	0.03808

IID standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Source: Author calulations from NielsenIQ Consumer Panel. Last regression excludes California.

We then analyze these unit pricing regulations at a higher level of detail. State regulations vary across two dimensions: Posting and Formatting. Table 3 shows the breakdown of states along these dimensions. First, states can opt to have unit price posting be voluntary (seven states) or mandatory (nine states). Second, states can specify how unit prices are formatted when they are displayed.[22] Formatting regulations specify features including minimum font sizes, background colors, and label positioning. With the exception of California, only states that mandate unit price posting have formatting requirements. Excluding California, regulations are naturally ordered: no regulation, voluntary posting, mandatory posting (no formatting requirements), and mandatory posting (with formatting requirements).

Columns (3) and (4) continue the earlier analysis, but leverage the stringency of the regulations. Column

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

Table 3: Unit Price Regulations by State

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

Note: Table reports whether unit price posting is mandatory or voluntary for retailers and whether or not there are strict formatting requirements on how unit prices should be displayed (minimum font size, color, etc.).
Source: State laws.

(3) shows that mandatory posting is associated with significantly higher bulk buying compared to no regulation, but states with voluntary requirements may have even higher rates of bulk buying. However, as Table 3 shows, California is an outlier in this regulatory environment because is the only state with the unique combination of voluntary posting and strict formatting requirements. Because of this, we exclude California and reestimate the regression. Column (4) reveals that California is the primary driver of this effect, and states with voluntary posting have only slightly higher bulk purchasing rates compared to states without regulation. On the other hand, mandatory unit price posting is associated with a 2.7–3.6 percentage point increase in bulk buying. The point estimates for bulk buying in states with strict formatting requirements are lower than those in states without formatting requirements, but these estimates are not significantly different from each other. This pattern supports the intuition that standardized unit price presentation reduces cognitive costs, increases the salience of unit prices, and facilitates comparisons for consumers.

This estimation provides strong evidence of a relationship between unit pricing regulations and bulk purchasing. However, there is a risk of measurement error which may bias our results towards zero. In particular, we use the presence or absence of unit price regulations as a proxy for whether or not stores display unit prices. In reality, this may underestimate the prevalence of unit pricing because stores in states without regulations may opt to post unit prices. For example, a national chain may adopt regional or chain-wide pricing policies and those policies will be influenced by the strictest policies the store must abide by. Therefore, the fact that stores in unregulated states may post unit prices will bias our results towards zero because a portion of the comparison group is actually being "treated" by seeing posted unit prices. Appendix Section A.5 shows that there are stronger effects of mandatory strict unit price regulation for local and regional chains than for national chains, but across all retailer sizes, there are significant increases in bulk buying associated with states with unit pricing regulations.

There is also a risk of selection bias since these regulations were primarily adopted in the Northeast and West Coast regions of the United States. To provide causal evidence, we examine about 13,000 households that move once during their tenure in the data. About 11% of these households move between regulatory regimes, while the remainder are either local moves or moves that maintain their current regulatory regime.

To estimate the effect of unit-price regulations on these movers, we use a difference-in-differences specification:

$$BulkShare_{it} = \alpha + \beta_1 Reg_{it} + \gamma X_{it} + \lambda_i + \lambda_t + \epsilon_{it}, \tag{5}$$

where the variables are the same as in Equation (4), but we control for household fixed effects (λ_i) and standard errors are clustered at the household level.[23,24]

Table 4 reports the results of estimating Equation (5). Columns (1) and (2) show that a household's bulk buying is about one percentage point higher when they are in a state with unit price regulations.

Table 4: Mover Event Study (All Products)

Model:	(1)	(2)
Variables		
Regulation	0.0108**	0.0115***
	(0.0044)	(0.0044)
\$25-\$50k		0.0006
		(0.0029)
\$50k-\$100k		0.0072**
		(0.0034)
>\$100k		0.0141***
		(0.0042)
Demographics	N	Y
Fixed-effects		
Household FEs	Yes	Yes
Panel Year FEs	Yes	Yes
Fit statistics		
Observations	116,127	116,127
Adjusted R^2	0.62313	0.62536

Clustered (Household FEs) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Source: Author calculations from NielsenIQ Consumer Panel.

Unit pricing regulations are relatively simple to implement for both policymakers and retailers. Retailers will bear some initial setup costs of redesigning their price labels, but ongoing costs will likely be similar to current menu costs that firms bear.[25] Adopting unit pricing policies (like those recommended by the National Conference on Weights and Measures) would encourage bulk buying while imposing few costs. These findings support the broader assertion that increasing price transparency allows households to choose products that deliver more value.

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

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

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

5 Model

The previous section shows that cognitive costs affect the bulk buying decision. To decompose the contribution of this factor, we embed it into a discrete choice model of the household's purchase decision. The ideal setting would include a homogeneous good where demand is uncorrelated with income. Because storage costs may also be a factor affecting product choice, we include measures of storage costs as well. Given substantial price, package size, and regulatory variation, differences in large and small purchases between households would identify storage costs and differences in buying between regulatory regimes would identify cognitive costs. This setting is approximated by one in which products have limited dimensions of differentiation and storage costs can be separately identified from demand.

A discrete choice model of toilet paper purchases closely approximates this ideal setting. Toilet paper is an excellent product for this analysis because it is a necessity item with easily observable dimensions of differentiation, namely price, quality, quantity, and package size. It is offered in a wide range of package sizes and stores stock numerous brands and sizes (grocery and mass merchandise stores usually stock 35–40 unique brand-sizes). The top five brands and private-label store brands account for 86% of sales. We focus on the most common package sizes, which range from 4- to 24-roll packages. We define a product as a unique brand-size combination.[26]

Additionally, underlying toilet paper consumption is primarily a function of household composition and age, not income.[27] High-income households consume a similar amount as low-income households but make fewer purchases (Orhun and Palazzolo, 2019). Finally, toilet paper cannot be easily substituted for another product nor can it be obtained through home production.[28]

The biggest identification challenge is separately identifying storage costs from underlying demand (i.e., households may buy large quantities because they have high consumption or because they have low storage costs). To separate storage costs from demand, we use variation induced by differences in product "concentration," which we define as the yield of the product per unit volume. Product concentration breaks the direct link between volume and consumption. In the detergent category, a product's yield is the number of washes it will supply. A concentrated detergent can wash the same number of loads but requires a smaller fluid volume than diluted detergent. Therefore, given the same number of washes, households that choose concentrated detergent must have higher storage costs than those choosing diluted detergent, assuming quality does not differ based on concentration.

The same reasoning holds true for toilet paper. Households do not demand a particular number of rolls (the primary determinant of package size), but choose how long they want their supply to last (i.e., purchase

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

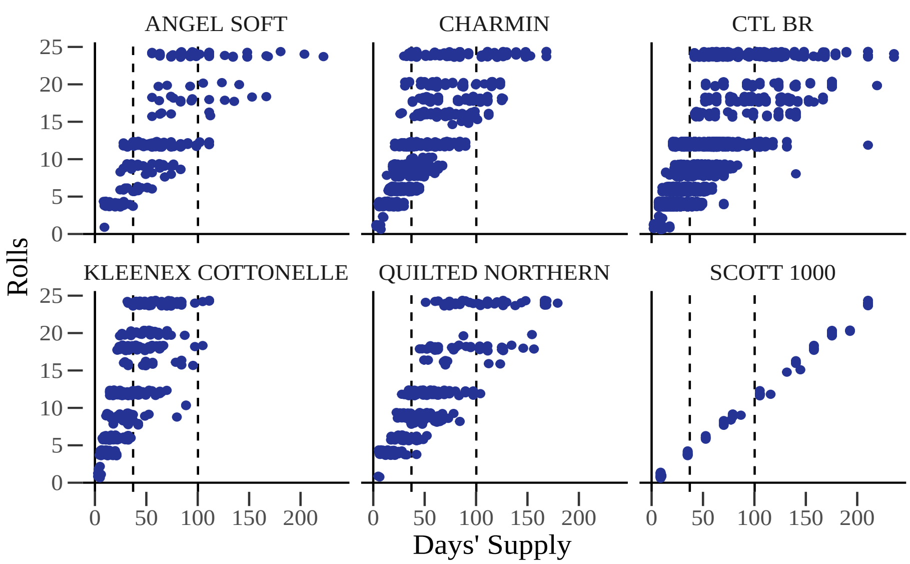
[27]A 100-fold cross-validated elastic net regression of annual purchases on household characteristics rules out income as significantly predictive. See Appendix A.6 for details.

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

enough to last for two weeks, a month, two months, etc.).[29] Toilet paper comes in a variety of concentrations with "mega" rolls being four times more concentrated than "regular" rolls. Therefore, a household that purchases 24 "regular" rolls has the same demand for toilet paper as a household that purchases six "mega" rolls, but the former household has lower storage costs since they can store the bigger package.

To illustrate the varying concentrations of toilet paper, Figure 4 plots the distribution of quantity (measured in number of days the supply will last for a single person) against package sizes (measured in rolls) for toilet paper products in the NielsenIQ data. As expected, there is an increasing relationship between how long the package will last and the number of rolls in a package, but there is substantial variation within packages containing the same number of rolls. The dashed lines denote the 25th and 75th percentiles of the average days' supply purchased by households. A wide range of package sizes fall within this range for each brand.[30] For example, a household demanding a 60-day supply of Charmin could purchase a package containing anywhere from 8 to 24 rolls. This overlap generates the necessary variation to separate storage costs from underlying demand.

Figure 4: Scatterplot of Toilet Paper Package Size and Quantity



Note: Figure plots the package sizes and quantities of the top five toilet paper brands and private-label products. The y-axis represents the number of toilet paper rolls contained in a package, while the x-axis represents the number of days a product will last a single-person household assuming a consumption rate of 57 two-ply sheets per day (Jaffe, 2007). Noise is added vertically to better illustrate the number of products available within package sizes since roll counts are discrete. Dashed lines indicate the 25th and 75th percentiles of the average days' supply purchased by households. **Source:** NielsenIQ Consumer Panel (2004–2019)

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

[30] Scott toilet paper is an exception because it does not offer different roll types. All rolls have 1,000 sheets.

5.1 Model Setup

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

$$\begin{aligned} U_{ijt} =& \beta_1 Price_{jt} + \beta_2^i UnitPrice_{jt} + \beta_3 UnitPrice_{jt} \times Reg_i + \\ & \beta_4^i \log(Days_j) + \beta_5^i BigPack_j + \beta_6 BigPack_j \times House_i + \\ & \beta_7^i SmallPack_j + \beta_8 SmallPack_j \times House_i + \theta_{b(j)} + \epsilon_{ijt}, \end{aligned} \tag{6}$$

where $Price_{jt}$ is the total price of product j at time t. Reg_i is an indicator for whether unit price regulations are in effect for household i. $Days_j$ is the number of days the package will last (a function of the number of total sheets in the package and the number of people in the household). $UnitPrice_{jt}$ is the per-day, per-person price of the package, since the yield of a package is how many days it will last. $BigPack_j$ is a dummy for the package having more than 12 rolls, and $SmallPack_j$ is a dummy for less than 12 rolls.[31] $House_i$ is an indicator for whether the household lives in a single-family home, with the alternative being an apartment. Finally, $\theta_{b(j)}$ is a brand fixed effect. Brand fixed effects capture quality differences between products. Because households may weigh unit prices or package sizes differently based on unobserved factors, we allow $\beta_2, \beta_4, \beta_5, \beta_7$ to vary. In particular, we assume they are normally distributed and allow for them to be correlated. We assume ϵ_{ijt} is iid Type 1 extreme value.

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

The price coefficient is identified using price variation across shopping trips due to shopping at different stores or sales. The size coefficient is identified by variation in the product "concentration" as illustrated in Figure 4. That is, given their preferred days' supply (x-value), some households choose large packages and some choose small packages (y-value).

Typically, researchers would be concerned about the endogeneity of prices and demand shocks. In our setting, we are not worried about this issue. Demand shocks seem unlikely as toilet paper is not a product that seems to undergo significant changes in underlying demand. An exception may be a major outbreak of a disease with substantial gastrointestinal symptoms. The main candidate for such an occurance would be influenza, but the 2018-19 influenza season was not particularly bad compared to others years.[32] Furthermore, while there was panic buying of toilet paper during the COVID-19 pandemic, the first COVID-19 case in the U.S. did not occur until January 2020 (Moore, 2020).

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

[32] Excluding the COVID-19 affected 2021-22 flu season, only one year since 2012-13 had significantly fewer flu cases than 2018-19 (https://www.cdc.gov/flu-burden/php/data-vis/past-seasons.html).

We estimate the model using simulated maximum likelihood.[33] To increase accuracy and reduce computational burden, we use pseudo-random Halton draws in the estimation procedure (Hensher and Greene, 2003).

6 Estimation and Counterfactual Results

We estimate this model separately for each income quartile using household purchases from 2019. We observe about 45,600 toilet paper purchases across about 14,400 households at grocery stores and mass merchandisers.

Table 5 reports model estimates for the random coefficients specification. The estimation results show that both the price and unit price coefficients are negative, implying that all else equal, households prefer lower prices. The interaction terms reveal that when unit prices are posted, all households are more sensitive to unit prices. This pattern supports the assertion that households respond to the provision of new price information. All households prefer to have more days' supply of toilet paper compared to less. In terms of storage costs, all households select against large sizes and, with the exception of the highest-income households, this preference is not significantly different based on their housing type. On the other hand, some households also dislike small packages.

Table 5: Random Coefficient Estimation Results (2019)

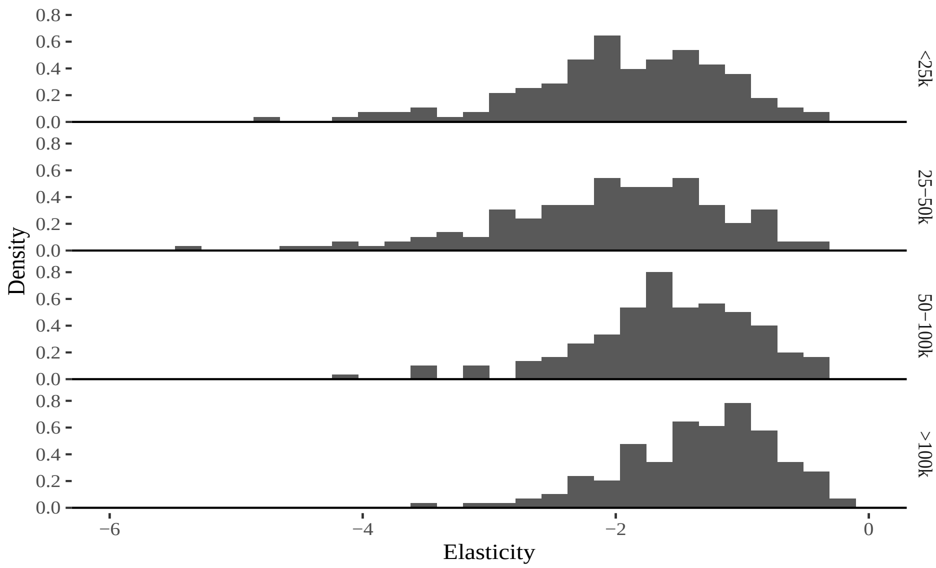
	<25k	25-50k	50-100k	>100k
	(1)	(2)	(3)	(4)
Total Price	-0.138^{***}	-0.179^{***}	-0.132^{***}	-0.120^{***}
	(0.011)	(0.007)	(0.005)	(0.007)
Unit Price	-1.965^{***}	-1.258^{***}	-0.837^{***}	-0.374^{***}
	(0.293)	(0.123)	(0.082)	(0.119)
. : Reg	-2.936^{***}	-0.987^{***}	-1.012^{***}	-0.682^{***}
	(0.319)	(0.168)	(0.100)	(0.110)
Log(Days)	0.489^{***}	0.927^{***}	0.788^{***}	1.048^{***}
	(0.089)	(0.053)	(0.043)	(0.067)
Large Size	-1.933^{***}	-1.398^{***}	-1.145^{***}	-0.750^{***}
	(0.240)	(0.122)	(0.102)	(0.152)
. : Home	0.329^{*}	0.249^{**}	0.105	-0.332^{**}
	(0.179)	(0.113)	(0.096)	(0.153)
Small Size	0.308^{***}	0.245^{***}	-0.264^{***}	-0.329^{***}
	(0.104)	(0.066)	(0.064)	(0.108)
. : Home	-0.783^{***}	-0.201^{***}	-0.135^{**}	-0.071
	(0.114)	(0.071)	(0.067)	(0.111)
Brand FE's	Y	Y	Y	Y
Observations	4,282	12,150	19,083	10,109
Log Likelihood	$-12{,}276.890$	$-35{,}470.040$	$-56{,}573.260$	$-29{,}920.070$
Note:	*p<0.1; **p<0.05; ***p<0.01			

[33] We use the "mlogit" package, which implements Ken Train's Matlab code in R (Revelt and Train, 1998; Croissant, 2020).

Each of the random coefficients displays substantial heterogeneity.[34] Overall, lower income households exhibit more heterogeneity in their sensitivity to unit prices, as implied by the standard deviations of the unit price coefficient.

Figure 5 plots the distribution of own-price elasticities for each product using the random coefficients estimates. The majority of elasticities fall between -1 and -4, with poorer households having larger elasticities (in magnitude).[35]

Figure 5: Distribution of Price Elasticity by Household Income



Note: Figure plots the distribution of price elasticities resulting from the estimation of Equation 6, using random coefficients.
Source: NielsenIQ Consumer Panel (2016) and NielsenIQ Retail Scanner (2016)

6.1 Model Fit

We examine model fit by comparing how well the model predicts the average size purchased for each income group. Since coefficients are random, the choice probabilities take the following form:

$$P_{ijt} = \int \frac{e^{\beta' x_{ijt}}}{\sum_j e^{\beta' x_{ijt}}} f(\beta) d\beta, \tag{7}$$

We use simulation to approximate the integral by taking 1,000 draws from the joint distribution of β. Table 6 compares the overall model predictions to the actual data. Overall, the fit is close, but the model overpredicts the amount purchased across all households, primarily because it overpredicts the purchases of particularly large generic packages. For example, a particular generic 12-pack has a 1%–2% share for each income group,

[34] Since each random coefficient was assumed to be normally distributed, some households may have a non-intuitive valuation for product attributes, such as a positive valuation for unit price. Sign restrictions can be imposed by assuming alternative distributions, such as a log-normal distribution.

[35] Table 4 of Cohen (2008) reports elasticities ranging from -1.94 to -2.54 for paper towels. Our estimates cover this range but have larger tails.

but based on its characteristics, the model predicts a 3%–5% share. The model assumes that all generic brands are equal, but in reality, it may be the case that generic brands differ based on the retailer that sells them. This additional dimension of heterogeneity could be captured by more granularly defining brands by the retailer that sells them.

Table 6: Random Coefficient Model Fit (Days' Supply Purchased)

Income	Data	Model
<25k	50.76	52.45
25-50k	50.54	52.02
50-100k	55.55	55.59
>100k	59.30	60.01

Note: Table compares the average days' supply of toilet paper purchased in the data with the predicted purchase from the model. We assume an average daily consumption rate of 57 two-ply sheets per day (Jaffe, 2007).
Source: NielsenIQ Consumer Panel (2019) and NielsenIQ Retail Scanner (2019)

6.2 Counterfactual

Using the parameter estimates from the previous section, we predict how households respond to universal unit price regulation. We compare all counterfactual results to a "base case" of predicted purchases given their current shopping environment.

We set each household's unit price coefficients equal to the sum of its coefficient and the regulation interaction term. For households making under $25,000, their unit price coefficient becomes $-1.965 - 2.936 = -4.901$.

Table 7 reports the counterfactual predictions for the random coefficients model.

Table 7: Bulk Purchasing Counterfactual Simulation Results

Income	Base	+ Unit Price Regs
<25k	51.03	52.67
25-50k	51.76	55.23
50-100k	53.07	55.91
>100k	55.94	55.83

Note: Table reports predicted package quantities purchased by households using model estimates of Equation (6). Units are the number of days the chosen package will last assuming average daily consumption rate of 57 two-ply sheets (Jaffe, 2007). "Unit Price Regs" imposes unit price regulations everywhere.
Source: Author calculations from NielsenIQ Consumer Panel.

The random coefficient counterfactuals, while overpredicting the average days' supply purchased, predicts a gap of 4.91 days' supply between high- and low-income households. After universally adopting unit price regulations, all households except the richest ones increase their purchasing and the gap between high- and low-income households shrinks to 3.16 days' supply. This counterfactual supports the main finding from Section 4, which showed that unit price regulations increase bulk buying.

7 Conclusion

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

By using state-level variation in whether or not retailers have to display unit prices, we find that displaying unit prices reduces cognitive costs and increases bulk buying. We then embed this factor into a discrete choice model of toilet paper purchases and predict how households' bulk purchasing changes if unit-price regulations are adopted universally and if storage costs are reduced. We find that posting unit prices closes the bulk buying gap by 36%.

This paper is one of the first to focus on consumer's take-up of quantity discounts and explore the factors that contribute to this decision. It provides evidence that cognitive costs affect a household's bulk buying decision. These differences have substantial financial consequences for the poorest households and are likely to generate systematic underestimates of consumption inequality if quantity discounts offset quality differences between products.

Furthermore, there is a growing awareness of a sneaky form of price adjustment known as "shrinkflation" where manufacturers and retailers keep the shelf price the same while reducing the quantity sold (Rosalsky, 2021, 2024). This kind of stealthy price adjustment is much harder for consumers to recognize, but posting unit prices will make it readily apparent to any consumer. Additionally, if the prices of large and small packages evolve differently, then households may experience substantial changes in their buying power. Future work will determine the extent to which inequality and inflation measures are misestimated because of quantity discounts.

References

Aguiar, M. and E. Hurst (2005). Consumption versus expenditure. Journal of Political Economy 113(5), 919–948.

Aguiar, M. and E. Hurst (2007). Life-cycle prices and production. American Economic Review 97(5), 1533–1559.

BLS (2019, April). Consumer expenditures in 2017. Technical Report 1080, Bureau of Labor Statistics.

Bogomolova, S. and I. Jarratt (2016). Unit pricing in supermarkets: Review of past evidence from academic and industry studies. Available at SSRN 2853977.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. American Economic Review 99(4), 1145–77.

Cohen, A. (2008). Package size and price discrimination in the paper towel market. International Journal of Industrial Organization 26(2), 502–516.

Croissant, Y. (2020). mlogit: Multinomial Logit Models. R package version 1.1-0, https://cran.r-project.org/package=mlogit.

Fellowes, M. (2006). From poverty, opportunity: Putting the market to work for lower income families.

General Accounting Office (1975). Food labeling: Goals, shortcomings, and proposed changes. Technical report, General Accounting Office.

Griffith, R., E. Leibtag, A. Leicester, and A. Nevo (2009). Consumer shopping behavior: How much do consumers save? Journal of Economic Perspectives 23(2), 99–120.

Hensher, D. A. and W. H. Greene (2003). The mixed logit model: the state of practice. Transportation 30(2), 133–176.

Jaffe, C. (2007). On a roll. MarketWatch.

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

Kunreuther, H. (1973). Why the poor may pay more for food: theoretical and empirical evidence. The Journal of Business 46(3), 368–383.

Mitchell, V.-W., D. Lennard, and P. McGoldrick (2003). Consumer awareness, understanding and usage of unit pricing. British Journal of Management 14(2), 173–187.

Miyazaki, A. D., D. E. Sprott, and K. C. Manning (2000). Unit prices on retail shelf labels: an assessment of information prominence. Journal of Retailing 76(1), 93–112.

Moore, A. (2020). How the coronavirus created a toilet paper shortage. NC State University College of Natural Resources News.

NIST (2019). Uniform laws and regulations in the areas of legal metrology and fuel quality (handbook 130). Technical report.

Orhun, A. Y. and M. Palazzolo (2019). Frugality is hard to afford. Journal of Marketing Research 56(1), 1–17.

Revelt, D. and K. Train (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. Review of Economics and Statistics 80(4), 647–657.

Rosalsky, G. (2021). Beware of 'shrinkflation,' inflation's devious cousin. NPR.

Rosalsky, G. (2024). How to fight shrinkflation? pay attention to unit prices at grocery stores. NPR.

Rose, V. (2000, November). Unit pricing and electronic shelf labels. Technical Report 2000-R-1044, Connecticut Office of Legislative Research.

Sprott, D. E., K. C. Manning, and A. D. Miyazaki (2003). Grocery price setting and quantity surcharges. Journal of Marketing 67(3), 34–46.

Statista (2019). U.s. population: Do you use toilet paper?

USDA (2019). Supplemental nutrition assistance program participation and costs. Technical report.

Appendices

A Appendix (For Online Publication)

A.1 Data Appendix

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

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

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

To construct our analysis sample, we remove any households in which the head of household is a student or a member of the military because these households likely have different living arrangements that are not representative of a typical household's decision (i.e., on campus housing or barracks are different than traditional homes and apartments). We drop any households living in mobile homes as well because this type of housing could include a wide range of house types including RVs and manufactured homes. We also remove any households making less that $5,000 and those that could not be geocoded based on their zip code.[36] Finally, some households were dropped because they could not be matched to tract-level vehicle access data.[37] Table A1 reports how many households were removed based on this cleaning procedure.

In the purchase data, we exclude alcohol, tobacco, pet items, health and beauty items, general merchandise, "magnet," and "deferred" product categories from our analysis. Alcohol and tobacco are excluded because of

[36]We use the 2017 Census Gazetteer to assign zip codes to the latitude and longitude of their population-weighted centroid.

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

Table A1: Homescan Sample

Step	HH
Starting HH:	$194,551$
Exclude military and students:	$191,149$
Exclude Households under 5k:	$189,994$
Exclude Mobile Homes:	$182,447$
Drop ZIPs Not Geocoded:	$181,694$
Drop ZIPs w/o Car Share:	$181,481$

their addictive qualities, which may induce peculiar purchase patterns. For example, a smoker may choose to only buy one pack of cigarettes with the intention of quitting even though a full carton may deliver a better value. Pet items are excluded to focus on products intended for human consumption. We exclude health and beauty items and general merchandise because these products such as trash cans, printers, eye shadow, and antacids are unlikely to be bought in bulk or have irregular consumption patterns. "Deferred" categories are categories that NielsenIQ has stopped tracking, so to maintain a consistent sample of products, these are excluded from our analysis. Finally, "magnet" purchases are items that do not have a UPC code such as fresh fruits and vegetables, deli counter items, or bakery items. Because these items are only recorded for a subset of NielsenIQ households and are not standardized, we also exclude them from our analysis. This process leaves us with 655 unique product categories. Overall, the products analyzed are common household staples including almost all food categories, basic toiletry items, and non-food essentials like toilet paper, soaps/detergents, and diapers.

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

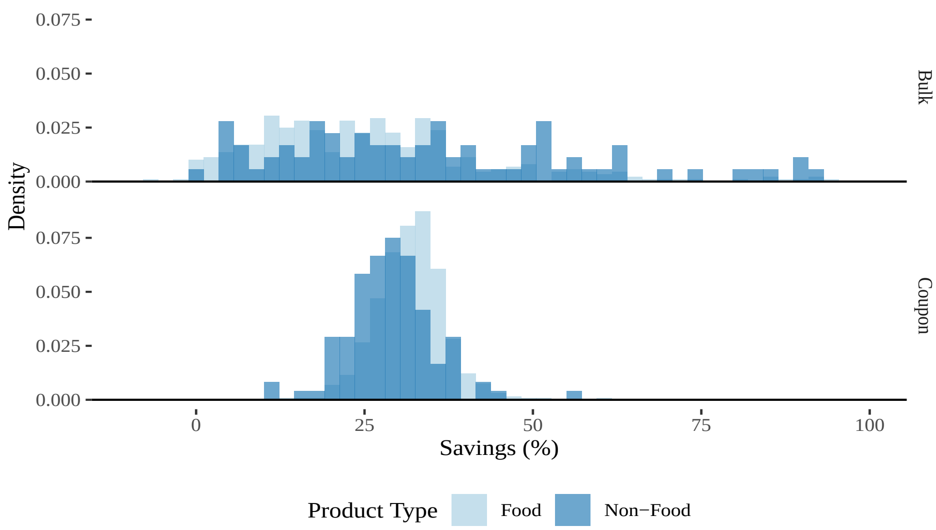
A.2 Quantity Discounts and Coupon Savings

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

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

Figure A1 plots the distribution of coupon savings and estimated bulk savings for food and non-food products. Coupon savings are narrowly clustered with a median savings of 29% for non-food products and 32% for food products. Bulk discounts have lower median savings for non-food and food products of 25% and 24%, respectively, but are more widely dispersed, even exceeding 75% savings for some non-food products.[38] Coupon savings are similar across product categories, while there is substantial variation in quantity discounts with non-food products offering higher savings.

Figure A1: Estimated Savings from Coupons and Bulk Discounts



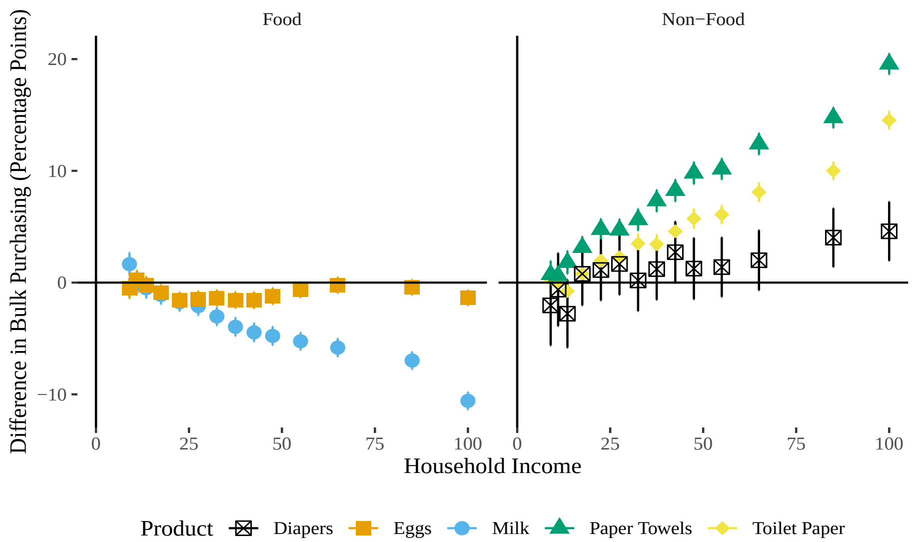
Note: Figure plots the distribution of savings from coupons and quantity discounts. For each coupon redemption, the percent savings are the ratio of the coupon value to the product's price. These savings are then averaged across all purchases in that product category. Bulk discounts are computed using coefficient estimates obtained from Equation (1) relating log unit prices to log package sizes. Bulk savings are the estimated savings obtained from moving from the second to the fourth quintile of the size distribution for each product category. **Source:** NielsenIQ Consumer Panel (2004–2019) and NielsenIQ Retail Scanner (2019)

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

A.3 Bulk Buying Across Popular Categories

Across popular spending categories, these gaps are particularly large in storable, non-food categories like paper towels and toilet paper, where households making over $100,000 are more than twice as likely to buy in bulk compared to households making under $25,000. In popular food categories like milk and eggs, there is little relationship or even a negative relationship between income and bulk buying (see Figure A2).

Figure A2: Bulk Purchasing by Household Income (Selected Products)



Note: Figure plots the income bin coefficients from Equation (2), which regresses the share of annual purchases that were bulk packages on household characteristics as well as market and year fixed effects. This regression is estimated for milk, eggs, diapers, toilet paper, and paper towels. NielsenIQ projection weights are used to ensure national representativeness. Households making $5k–$8k are the reference group. Standard errors are clustered at the county level. **Source:** NielsenIQ Consumer Panel (2004–2019)

A.4 Alternative Calculation of Missed Quantity Discounts

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

We compute the first-best savings a household could obtain for toilet paper, diapers, milk, and eggs using the following approach. First, for each shopping trip, we compute the lowest unit price the household could have paid given its brand and store choice in that week. The difference in unit prices relative to the unit price chosen is a household's first-best savings for that purchase. Then, to get the average savings for a household, we compute the expenditure-weighted average savings across all purchases for each household. Based on this measure, Table A2 reports average excess spending by income group, computed for a family of four.

Table A2: First-Best Savings by Income Group

Income Group	Savings
<25k	0.36
25-50k	0.35
50-100k	0.34
>100k	0.33

Note: Table shows average first-best savings for a family of four.
Source: Author calculation based on NielsenIQ data.

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

$$Y_{imt} = \alpha + \sum_{q} \beta^q Income_{imt} + \gamma X_{imt} + \lambda_{mt} + \epsilon_{imt}, \tag{8}$$

where Y_{imt} is the excess spending of household i in market m in year t. $Income_{imt}$ is the household's income bin and X_{imt} consists of household characteristics. λ_{mt} is a market-year fixed effect. Table A3 shows that in non-food categories low-income households miss out on 1–2 percentage points more savings than high-income households while this effect disappears for food categories. Given the perishability of food items, these savings may not be realized if the product perishes before it can be consumed.

Table A3

	Diapers	Toilet Paper	Eggs	Milk
	(1)	(2)	(3)	(4)
25-50k	-0.008^{**}	-0.004^{***}	0.004^{***}	0.009^{***}
	(0.004)	(0.001)	(0.001)	(0.001)
50-100k	-0.015^{***}	-0.007^{***}	0.009^{***}	0.019^{***}
	(0.004)	(0.001)	(0.001)	(0.001)
>100k	-0.019^{***}	-0.010^{***}	0.019^{***}	0.032^{***}
	(0.004)	(0.001)	(0.001)	(0.001)
Demographics	Y	Y	Y	Y
Market-Year FE	Y	Y	Y	Y
Observations	50,632	318,404	399,974	462,805
Adjusted R^2	0.031	0.062	0.107	0.123

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1
Source: Author calculations from NielsenIQ Consumer Panel. Dependent variable is share of bulk purchases. Standard errors are clustered by state.

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

A.5 Bulk Buying by Store Type or Chain Size

In this section, we analyze whether the effect of unit pricing differs by store type or chain size. Unit price regulations are only at the state level, but retailers are free to post (or not post) unit prices as long as they are within the boundaries of the law. Large chains may post prices uniformly across all stores in a way that meets the strictest requirements they are subject to. On the other hand, regional chains or independent stores may more closely mirror the laws of the state they are located in. We estimate Equation 4 using annual household bulk buying at specific stores types or within different chain sizes. Each observation is at the household-year-channel (or chain) level. For example, bulk items accounted for 50% of Household A's grocery store purchases, while bulk items accounted for 100% of Household A's warehouse club purchases.

Table A4: Correlation of Bulk Buying and Unit Price Regulation by Channel Type

	Grocery	Drug Store	Discount Store	Dollar Store	Warehouse Club
Voluntary	0.0042	-0.0128***	-0.0079	-0.0076*	-0.0041*
	(0.0090)	(0.0035)	(0.0049)	(0.0044)	(0.0021)
Mandatory	0.0276***	0.0105	0.0123*	-0.0180**	0.0084***
	(0.0047)	(0.0063)	(0.0073)	(0.0080)	(0.0028)
Mandatory Strict	0.0525***	0.0198***	0.0040	-0.0037	0.0042**
	(0.0087)	(0.0031)	(0.0036)	(0.0056)	(0.0020)
Avg Bulk	0.37	0.29	0.49	0.35	0.95
Demographics	Y	Y	Y	Y	Y
Omit CA	Y	Y	Y	Y	Y
Observations	710,832	332,118	648,396	380,222	312,127
Adjusted R^2	0.01140	0.00296	0.00615	0.00256	0.00116

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Source: Author calculations from NielsenIQ Consumer Panel. Excludes California. Dependent variable is share of bulk purchases. Standard errors are clustered by state.

Table A4 shows that in the cross-section, stricter unit price regulations are associated with more bulk buying primarily for grocery stores, drugstores, and warehouse clubs. Households in states with strict unit price regulations buy in bulk five percentage points more at grocery stores compared to households in states without any pricing regulations. Since grocery stores tend to be regional or independent, the large positive relationship provides strong evidence that unit price regulations can increase bulk purchasing. Grocery stores also have the richest variety in NielsenIQ's data with over 900 unique retailers being captured compared to 66 drugstores, 27 discount stores, 18 dollar stores, and 10 warehouse clubs.[39] Other store types exhibit smaller or insignificant effects, which could be because these are generally large chains that have more uniform pricing practices across all locations.

Table A5 shows the results by chain size. Following Jarmin et al. (2009), we define a "local" chain as only

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

Table A5: Correlation of Bulk Buying and Demographics by Chain Type

	Local	Regional	National
Voluntary	0.0813	0.0339	0.0049
	(0.0722)	(0.0266)	(0.0094)
Mandatory	-0.0435	-0.0176	0.0375***
	(0.0635)	(0.0126)	(0.0085)
Mandatory Strict	0.1080***	0.0814***	0.0259***
	(0.0195)	(0.0120)	(0.0060)
Avg Bulk	0.3400	0.3300	0.5000
Demographics	Y	Y	Y
Omit CA	Y	Y	Y
Observations	918	45,112	773,091
Adjusted R^2	0.02681	0.00555	0.03792

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Source: Author calculations from NielsenIQ Consumer Panel. Excludes California. Dependent variable is share of bulk purchases. Standard errors are clustered by state. "Local", "Regional", and "National" refer to chains with locations in one state, in two to ten states, and in more than ten states, respectively.

having locations in one state, a “regional” chain has locations in two to ten states, and a “national” chain has locations in more than ten states. In the cross-section, strict unit price regulations are associated with more bulk buying across all chain types. The effect is strongest for local and regional chains, exhibiting a eight to ten percentage point increase in bulk buying relative to states without unit price regulations. National chains still have significant differences, but they are a more moderate two to four percentage point difference relative to states without regulations. Overall, the relative effect is strongest for the smaller chains that are likely to only be subject to a limited set of regulations and the effect is weaker for national chains that may be more likely to adopt pricing practices that satisfy the strictest requirements nationwide.

A.6 Annual Consumption Analysis

We show that income is not predictive of a household’s toilet paper consumption rate. If income and toilet paper consumption are related, then an OLS regression will extract the correlation.

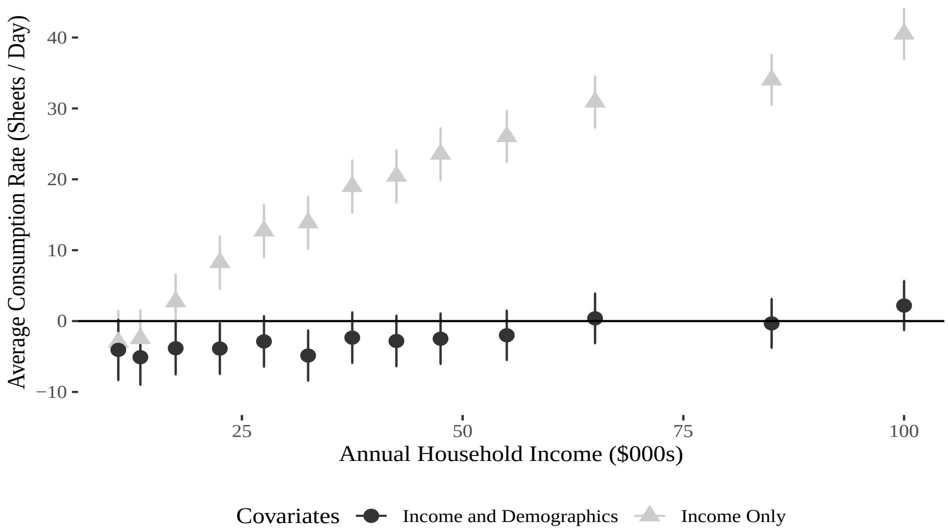
First, we compute a household’s daily consumption by aggregating the total number of sheets purchased by a household in a given year, excluding the final purchase of the year since it may not be consumed within the year. We divide this total by the number of days between the first and last purchase of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

Given a household’s average daily consumption rate, we estimate an OLS regression of consumption on household characteristics:

$$Y_i = \alpha + \beta X_i + \epsilon_i, \tag{9}$$

where Y_i is household i's average daily consumption and X_i is a vector of household characteristics. Figure A3 plots the income coefficients of an OLS regression including only income covariates and the coefficients when household characteristics are included. The graph illustrates that after controlling for covariates that plausibly cause increased consumption, income is not significantly correlated with consumption.

Figure A3: Average Daily Consumption by Household Income



Note: Figure plots the income bin coefficients from Equation (9), which regresses average daily household toilet paper consumption on household characteristics. Average daily consumption is computed by dividing total quantity purchased in a year by the number of days a household was active in the panel. **Source:** NielsenIQ Consumer Panel (2004–2019)

Table A6: NielsenIQ Consumer Panel Summary Statistics by Unit Price Regulation

	Without Regs		With Regs	
Variable	Mean	SD	Mean	SD
Household income ($000s)	56.63	30.98	60.30	31.87
Household size	2.53	1.43	2.61	1.49
Child Present	0.32	0.47	0.32	0.47
Married	0.52	0.50	0.49	0.50
College Educated	0.38	0.49	0.42	0.49
Age	52.45	14.42	53.14	14.45
N (Household-Years)	559,185		290,878	

Note: Unweighted means and standard deviations are reported.

Source: Own calculation based on NielsenIQ Consumer Panel (2004–2019)