Online Appendix

Shrinkflation and Consumer Demand*

Aljoscha Janssen[†] Johannes Kasinger[‡]

September 1, 2025

A Data

In the following sections, we provide a detailed explanation of our approach. We begin by explaining the general brand selection process and the identification of size changes at the store level. We then outline the additional data manipulation steps that preceded the empirical model of consumer response, including retailer selection and price imputations.

A.1 Brand Selection

The initial data selection is based on brand sales. The identification of product size changes in section 2.2 at the store level requires sufficient sales volume for each product. To ensure this, we restrict our dataset to leading brands within each of the 1,100 product modules. We define "leading brands" as those that account for 80% of total sales revenue in their respective modules. Specifically, we calculate the aggregate sales over the 10-year period for each brand within each module and then consider only those brands within the top 80th percentile. This restriction allows for a more accurate estimation of retailer stock assortments on a weekly basis. The exclusion of non-leading brands forms the basis for identifying product size changes, as well as the summary statistics, descriptive insights, and demand analysis.

*Researcher(s)' own analyses calculated (or 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 researcher(s) 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.

[†]Singapore Management University, ajanssen@smu.edu.sg

[‡]Tilburg School of Economics and Management and Leibniz Institute for Financial Research SAFE, j.kasinger@tilburguniversity.edu

A.2 Product Size Changes

The primary challenge in our empirical exercise is identifying product size changes at the individual store level. Therefore, we provide a step-by-step guide for identifying candidates for size changes, along with measures to ensure that the changes are genuine and not merely alterations in assortment, branding, or design.

A.2.1. General equivalence. Before proceeding to the store level, we ensure that any candidate for a product size change is identical in brand, brand description, and UPC description but differs in actual UPC code and size. Specifically, we use the product description file (`products.tsv`), compare all products, and ensure that the variables `product_module_code`, `product_group_code`, `brand_code_uc`, `upc_descr`, and `multi` are identical but that `upc_code` differs. The size variables `size1_amount` must fall within a 25% range in pairwise comparisons. If multiple products meet these criteria, we include all in the equivalence set, allowing more than two products to be grouped as similar.[1]

We note three clarifications. First, a change in UPC can also occur if the number of individual items in a multipack changes (variable `multi`), for example, from a six-pack to a five-pack of soda. While this situation constitutes a size change, we do not consider these products identical and thus exclude them from our analysis of shrinking or increasing sizes. Second, occasionally, UPCs are reused for entirely different products. NielsenIQ's retail scanner addresses this with a yearly UPC version code (`upc_ver_uc`), mapping products in the product file to their correct codes in weekly store-level observations. This process does not affect our main results, as we only consider different UPCs as candidates for size changes. Third, in principle, a product with the same UPC might differ only in size. In such cases, size changes would be identifiable solely through `upc_ver_uc`, but because the version code updates annually, these changes are not detectable at specific points in time. Therefore, we exclude these cases. We provide a detailed analysis of this restriction in Online Appendix B.

A.2.2. Requirements at the store level. Having established a product overview of potential size-changing candidates, we now turn to the store level, in which our goal is to identify substitution patterns of products undergoing size changes. To ensure these are actual size changes and not changes in product assortment, we proceed as follows: within each module, we merge the product overview with NielsenIQ's `Movement` files.[2] We then assess whether the store-level

[1] We convert all size units to ounces when possible. In rare cases, when a product changes both UPC and size units to an incomparable unit, such as from ounces to count, we do not consider this a size change under our definition.

[2] This step requires accounting for the correct UPC version before the annual `rms_vers` files.

data indicate a size increase or decrease according to four conditions: First, we confirm that the size change occurs at the store level based on the absolute size difference of the products. Second, we exclude products outside the dataset's temporal boundaries (before March 31, 2010, or after October 1, 2020). Third, for a size change to be considered, the old product must exit the store when the new product arrives. The substituted product must not reappear within that store until the final week of the sample, which ends in December 2020. We allow for a maximum of eight weeks during which neither product is observed or both are present. Finally, we require both product versions to be available in a store for at least four weeks. Fourth, a few multiple-size changes are observable. In those cases, the product changes occur multiple times, such that we observe more than two UPCs that satisfy the general equivalence outlined previously. In such cases, a pairwise comparison described in the first three conditions is challenging. We adopt a conservative approach: (1) for each unique size, we include only products that are available the longest in a store over the observation period, and (2) from the remaining products, we select the two with the latest recorded sales in a store. This method ensures that we do not overestimate product size changes by avoiding repeated counting.

After applying this procedure, we are left with the instances we define as size changes at the store level. Online Appendix C provides further robustness checks.

A.3 Data Restriction for Demand Analysis

The price imputation described in Appendix A.4 and the demand estimation in Section 4 present significant computational challenges, particularly regarding memory usage. To ensure feasibility, we apply two key restrictions to the data: (1) We limit the sample to food, mass merchandise, and drug stores; (2) we include only UPC-UPC version combinations that experienced a size change in at least one store in the U.S., according to our definition, during the observation period (2010–2020). The second restriction ensures that all recorded weekly sales of original-sized and newly sized products are retained, even for stores where no size change occurred and for all periods before or after the size change. This approach provides a comparable set of products, keeps the estimation computationally feasible, and retains a wide range of observations unrelated to size changes, which allows for meaningful time-fixed effects. Due to computational constraints, we further treat the product modules "CANDY - CHOCOLATE" and "CANDY - NON-CHOCOLATE," both within product group "CANDY," and "SNACKS - POTATO CHIPS," within product group "SNACKS," as separate product groups. The estimates for product groups "CANDY" and "SNACKS" include all other product modules within these groups, excluding the separately estimated product modules.

A.4 Price Imputation

In the RMS retail scanner data, price information is only available for the weeks a product is sold. This leads to missing price observations, especially for smaller products with infrequent sales. We use the price imputation algorithm developed by Hitsch et al. (2021) to overcome this issue. The algorithm works by identifying both base and promoted prices. It does so by examining the typical price pattern at the store level, in which base prices are stable for some time, followed by shorter periods of promoted prices. The algorithm can fill in missing price data, providing a more accurate representation of the full price spectrum for these products. It assumes that weeks with no sales are periods when prices are not promoted.[3] The algorithm imputes prices if, for the last observed (base) price (i.e., the reference price), there is another observed price (close enough to the reference price) within a 12-week window. If this is the case, prices that are missing between these two observed prices will be set equal to the reference price, assuming that the reference price is correct and stable within the specified window. As a result, the price in a respective period can still be missing if there is no observed price within 12 weeks from the last observed price (for more details, see the online appendix of Hitsch et al., 2021).

Operationally, we process the data by loading the yearly movement files for each module and merging them with the corresponding annual UPC version and store files. Next, we filter the movement files based on the criteria outlined in Section A.3. After processing, we append all the yearly movement files within a module into a single dataset. Finally, we apply the imputation algorithm to this dataset, using the combination of the UPC and the respective UPC version code as the product identifier. For the seven largest modules, we first process and append the movement files within each module by store type to ensure computational feasibility.

Columns 4 and 5 in Table A.1 compare aggregated summary statistics of the original movement files with those of the imputed dataset. Through the price imputation algorithm, we increase the number of observations from nearly 14.5 billion to 22.1 billion observations, implying that approximately 34% of the observations in the final data set are imputed, a ratio that is similar to that of Hitsch et al. (2021). Unsurprisingly, the average units and value of purchase are both lower in the imputed dataset, as it now includes zero quantity observations. Similarly, the

[3] For demand estimation, we add 0.5 to the units sold before taking the logarithm to handle weeks with zero sales.

average price is slightly higher in the imputed dataset, as lower prices are positively associated with quantity. Average sizes slightly increase.

A.5 Summary statistics on data exclusion

Table A.1 provides summary statistics for various configurations of the raw dataset, showing how the dataset changes as we impose additional restrictions. We calculate descriptive statistics in each column using the underlying dataset of the preceding column imposing one additional restriction. The table displays summary statistics (1) for the full sample; (2) for the top 80th-percentile brands within each product module; (3) when only considering products with size change; (4) when only considering food, drugstores, and mass merchandise; and (5) for the final imputed data used for our regression analysis.

B Shrinking a Product without Changing the UPC

We primarily identify a product that is decreasing in size at the store level using the following criterion: a new UPC from the same brand enters the system. This new UPC, while having an identical brand and description to that of the previous UPC, is up to 25% smaller. Concurrently, the old UPC exits the system. However, our data source does not enable us to detect product shrinkage when a smaller product is substituted under the same UPC code. With product details for a specific UPC only updated at the start of the year, identifying the exact time of substitution for products with unchanged UPC codes is not feasible.

In this section, we contend that neglecting size reductions of same UPC products does not compromise the integrity of our analysis for two primary reasons. The first pertains to the prevalent practices of retailers and manufacturers. Most adhere to guidelines set forth by the Uniform Code Council, now recognized as GS1 US, the originators of the UPC system. The UPC is predominantly designed to identify individual consumer products at retail point of sale, serving as a distinct barcode predominantly used in North America for retail commodities.[4] According to the management standards stipulated by GS1, any alteration in net content mandates the introduction of a new GTIN and, consequently, a new UPC (GS1US, 2023). To verify the adherence of manufacturers to this practice, we explored popular consumer forums such as Reddit (`https://www.reddit.com/r/shrinkflation`), on which consumers voluntarily report

[4]The UPC is frequently used interchangeably with the GTIN (Global Shipment Identification Number). The GTIN varies in length, but the UPC-A barcode, the version most prevalent at points of sale and in our dataset, employs data from the GTIN-12 (GS1US, 2023).

recent instances of product shrinkage. In the majority of cases in which UPC information was accessible, we observed modifications to the UPCs. However, note that no stringent legal mandate enforces UPC alterations for every size change, which implies that some manufacturers might retain an unchanged UPC despite modifying product dimensions.

Second, we approach the question from an empirical perspective. While our current methods do not allow us to pinpoint instances in which products under the same UPC experience size reductions (given our inability to temporally and geographically trace the introduction of an identical UPC at the store level), we can juxtapose the overall count of such UPCs exhibiting varied sizes against the tally of multiple UPCs manifesting similar size variations nationwide from 2010 to 2020. By contrasting the aggregate of products with unchanged versus altered UPCs that display size differences within a 25% range, we gain preliminary insight into the frequency of occurrences in which products retain their original UPC despite size adjustments.

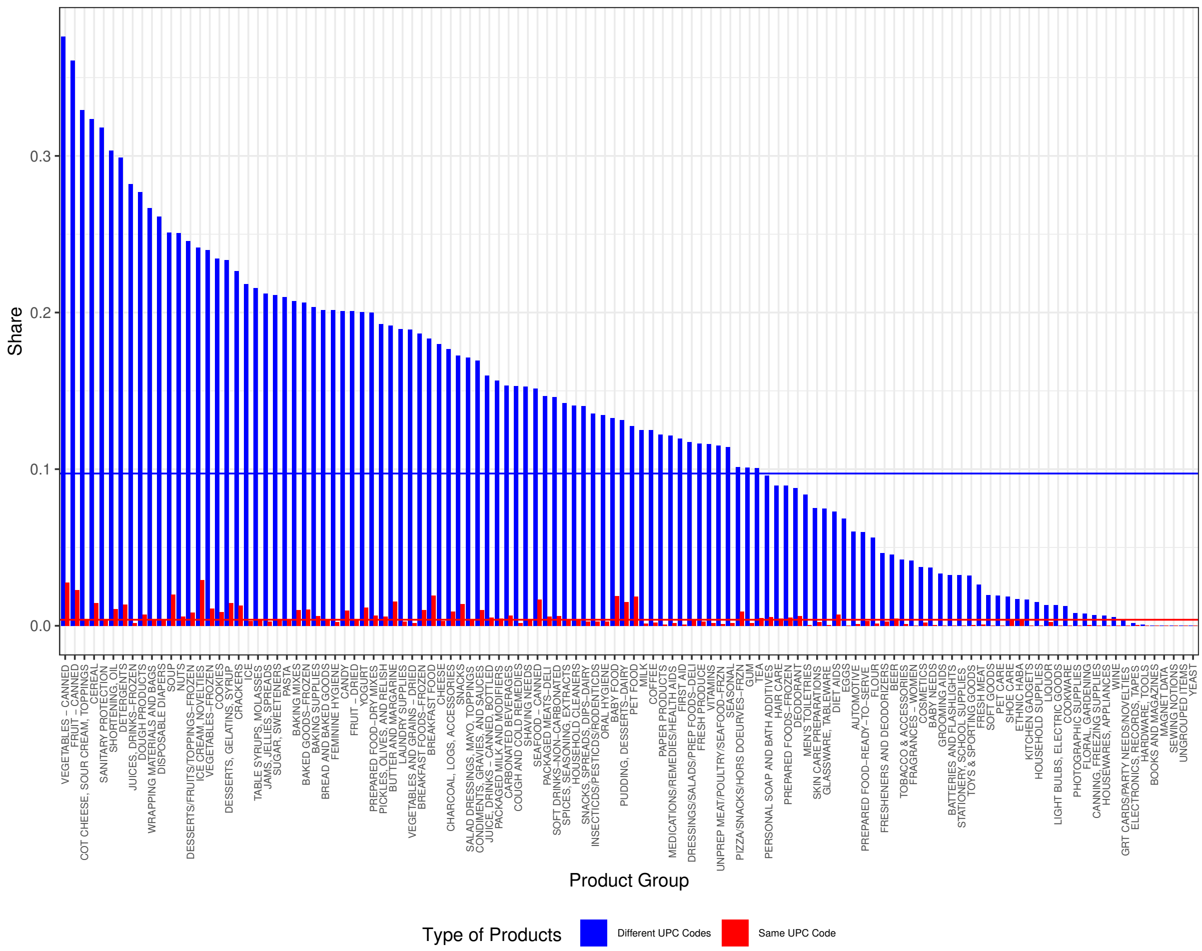
In Figure B.1, we present the proportion of products within each category that exhibit multiple size variations within a 25% range, while maintaining identical brand, brand description, and UPC description. We distinguish between products with differing UPCs (depicted in blue) and those with identical UPCs (depicted in red). On average, 9.7% of products display size variations alongside different UPCs, despite being marketed under the same brand with identical descriptions. In contrast, a mere 0.4% of products show size differences under the same UPC with the same brand and description. This marked discrepancy suggests that instances of product size changes without corresponding UPC updates are relatively rare. It also reinforces our belief that the vast majority of manufacturers and retailers likely adhere to GS1's management standards when implementing product changes such as size reductions or substitutions.

C Are Products Identical?

C.1 Visual Inspection

In our primary analysis, we focus on identifying products that exhibit decreasing or increasing trends. To achieve this, we examine products from the same brands and analyze UPC descriptions that are substituted at the store level with either a smaller or a larger equivalent. However, this approach is not without risks, particularly in terms of potentially mislabeling certain products. For example, a product could be replaced by another version that is not only different in size but also distinct in its marketing approach. Consider a scenario in which a product under-

Figure B.1: Different Sized Products, Same and Different UPCs



Notes: This graph illustrates the proportion of products that feature size variations within a 25% range, while retaining identical brand attributes and UPC descriptions. Each bar represents a distinct product category. The data are differentiated by products with variant UPCs (in blue) and those with consistent UPCs (in red). The solid blue and red lines indicate the average percentage of products with multiple sizes across all categories, corresponding to different and identical UPC scenarios, respectively.

goes a rebranding or packaging overhaul. Although such concurrent changes may be common, we acknowledge this limitation and address it by implementing several robustness checks.

In a first robustness check, we examine each individual UPC associated with products that have demonstrated either a decrease or an increase in size. This involves a detailed comparison of packaging to ascertain whether the products in question are truly identical. Given the sheer volume of products across all modules, manual verification for each is impractical. Therefore, we chose to concentrate on specific modules in which changes in product size are most commonly observed: candy, snacks, detergents, hair care, and cereals.

For each product within these modules, we consult a UPC lookup database (accessible at https://www.upcitemdb.com/) to accurately determine the packaging of each product both before and after any changes in size. However, this database has its limitations, particularly for products that were phased out more than a few years ago and are no longer on the market. Consequently, we rely on a subset of products with confirmed size changes as a representative sample to demonstrate the validity of our analysis. The remaining products may still be valid product size changes, but we are not able to confirm the changes visually.

In this analysis, we present a detailed comparison between the complete collection of products that have undergone size reductions and products specifically verified as identical despite this decrease. We detail the relevant statistics for the product group of snacks in columns 1 and 2 in Table C.1. We also show corresponding statistics for the product groups of hair care, detergent, and candy in Table C.1. Our focus centers on key aspects such as variations in product size, price dynamics, and sales trends before and after the size reduction event.

Across all four categories, we examine 4,270 unique UPCs and confirm that approximately 37.3% correspond to identical products. Unconfirmed cases do not necessarily indicate inaccuracies or different products; they often stem from insufficient information—typically missing images—in the UPC lookup database.

At the store-product level, where each observation represents a single instance of size reduction in a store, we can confirm identical products for 59.6% of cases (1,761,014 out of 2,955,739 observations). This higher share reflects the greater likelihood that common products appear in the database. Overall, the analysis suggests that a substantial share of observed size reductions involves the same product, reaffirming our approach.

When analyzing both sets of products against our selected metrics, only slight differences emerge. Crucially, trends in size reduction, price adjustments, and sales fluctuations are remarkably similar across the general range of products and those where the identity postreduction has been confirmed. This uniformity leads us to conclude that, apart from the change in size, the products remain essentially identical in other respects. This insight offers a valuable perspective on the market dynamics associated with changes in product size.

C.2 Household Data

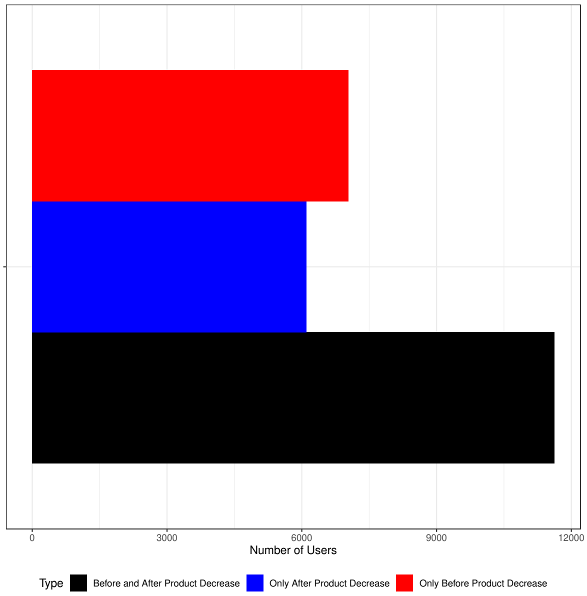
In our second robustness check, we employ NielsenIQ's household scanner data to examine whether households consistently purchase identical products before and after a change in package size. Despite the possible variations in consumer behavior, this pattern generally supports the accurate identification of products undergoing size changes. However, it is important to note that the NielsenIQ household scanner data are not perfectly suitable for this analysis. Product size changes at the store level are infrequently observed, and their occurrence is relatively rare compared with the number of households in the panel purchasing products within a specific segment. Consequently, our analysis is conducted on a small sample of observed households within a specific store.

We investigate each product size reduction recorded over the 10 years within our sample, focusing on households that purchased a product that decreased in size at least four times in independent shopping trips within one year before and one year after the size reduction. This criterion ensures that the analysis includes only households that regularly purchase the product, rather than those trying it incidentally without awareness of previous price changes, thereby precluding their purchase decisions from being influenced by size alterations.

Figure C.1 presents the results, depicting the number of households that exclusively purchased the product before, after, or both before and after the product size decrease. The data reveal that repeated purchases both before and after the size reduction are the most common. The frequencies of households that exclusively purchase the product either before or after the size decrease are relatively similar, with a marginal tendency toward more purchases before the size reduction.

While this analysis does not provide in-depth insights into the household choices regarding product size changes, due to infrequent purchases and the limited number of households buying specific products that decrease in size, the summary statistics suggest a higher likelihood of households purchasing the product both before and after a size increase. This observation further corroborates our belief that the products undergoing size decreases are identical in all dimensions other than size.

Figure C.1: Repeated Purchases by Households



Note: This graph shows the number of households purchasing products that decrease in size. Specifically, it includes households with at least four purchase occasions of products experiencing size decreases in the year before and the year after the event of a decrease. The graph categorizes the households into three groups: those that only purchased the product before the size decrease (red), after the size decrease (blue), and both before and after the size decrease (black).

Table A.1: Summary Statistics by Restrictions

	(1) Full data	(2) Top 80%-pctl. brands	(3) Food, Mass merch. & Drug stores	(4) Shrunk/increased UPCs	(5) Imputed Data
Total Observations	143.99 Bn	103.84 Bn	101.19 Bn	14.54 Bn	22.09 Bn
Unique UPCs	1,552,445	540,182	533,239	24,377	24,572
Unique Stores	59,352	59,345	44,537	44,534	44,534
Unique Retailers	141	141	121	121	121
Unique UPC-Store	2,551.12 Mn	1,554.92 Mn	1,514.60 Mn	158.74 Mn	160.27 Mn
Total Sales	2,643.62 Bn	2,044.99 Bn	1,974.14 Bn	306.68 Bn	305.91 Bn
Total Units	818.76 Bn	635.40 Bn	616.87 Bn	112.60 Bn	112.23 Bn
Avg. Price	4.67	4.71	4.72	3.97	4.44
Pooled SD Price	4.12	4.10	4.01	2.62	2.88
Avg. Product Size	29.89	30.34	30.09	28.36	29.49
Pooled SD Product Size	51.39	53.54	52.64	30.90	32.70

Notes: Column 1 shows statistics for the entire dataset. Column 2 restricts the dataset to brands comprising at least 80% of sales within a product module. Column 3 includes only food stores, drugstores, and mass merchandisers. Column 4 further limits the data to UPCs that underwent a product size change in at least one store during our study period, as defined in our methodology. Column 5 shows the statistics for the imputed data set based on the dataset in column 4. Average prices and sizes are calculated annually by product module and averaged across years and modules, weighted by the respective shares of total observations. Similarly, standard deviations are computed annually by product module and averaged across years and modules using the formulas: $SD_p = \sqrt{\sum_t (n_t - 1) \times SD_t^2 / \sum_t (n_t - 1)}$ *and* $SD_{all} = \sqrt{\sum_p (n_p - 1) \times SD_p^2 / \sum_p (n_p - 1)}$.

Table C.1: Comparison of Products That Are Confirmed by Visual Inspection in Different Product Groups

	Snacks		Candy		Detergents		Hair Care	
	All	Confirmed	All	Confirmed	All	Confirmed	All	Confirmed
Number of Observations	1,317,016	778,788	896,597	449,923	382,603	232,285	359,523	300,018
Number of Unique UPCs	1,385	590	1,893	512	541	253	451	237
Avg Size Before Decrease	7.61	7.72	13.08	14.72	39.70	35.58	16.38	16.30
	(4.04)	(3.75)	(10.28)	(11.51)	(34.17)	(31.68)	(7.33)	(7.22)
Avg Size After Decrease	7.05	7.19	11.79	13.35	34.79	31.20	14.52	14.42
	(3.76)	(3.53)	(9.24)	(10.28)	(29.93)	(27.16)	(6.26)	(6.20)
Avg Price Before Decrease	2.58	2.63	3.78	4.60	4.82	4.38	4.07	3.99
	(1.13)	(1.17)	(2.88)	(3.12)	(3.98)	(3.51)	(1.58)	(1.41)
Avg Price After Decrease	2.57	2.61	3.84	4.69	4.64	4.21	4.21	4.17
	(1.13)	(1.16)	(2.92)	(3.14)	(3.77)	(3.30)	(1.66)	(1.61)
Avg Price per Volume Before Decrease	3.76	3.86	5.25	4.70	13.25	13.74	6.30	6.20
	(6.39)	(6.57)	(11.16)	(10.81)	(35.65)	(37.87)	(15.82)	(15.27)
Avg Price per Volume After Decrease	3.43	3.51	4.30	3.60	11.58	11.72	7.61	7.26
	(5.65)	(5.82)	(7.41)	(6.05)	(27.95)	(28.32)	(21.99)	(21.42)
Avg Weekly Units in Store Sold Before Decrease	9.55	9.81	3.76	3.63	4.80	5.38	1.79	1.79
	(13.59)	(14.22)	(5.56)	(5.93)	(6.21)	(6.96)	(1.20)	(1.15)
Avg Weekly Units in Store Sold After Decrease	9.62	9.92	3.68	3.57	4.65	5.22	1.76	1.75
	(13.71)	(14.24)	(5.29)	(5.70)	(5.87)	(6.53)	(1.24)	(1.23)
Avg Weekly Sales in Store Sold Before Decrease	22.70	23.65	10.87	12.02	16.14	16.62	7.00	6.80
	(33.07)	(35.45)	(15.00)	(12.60)	(30.93)	(19.97)	(6.88)	(5.59)
Avg Weekly Sales in Store Sold After Decrease	23.11	24.02	10.98	12.26	15.54	15.77	7.13	6.94
	(33.97)	(35.98)	(14.49)	(12.61)	(24.97)	(17.81)	(7.08)	(6.28)

Notes: This table presents a comparative analysis of products within various subsegments (snacks, candy, detergents, hair care), focusing on those that underwent a size reduction. It differentiates between products confirmed to be identical despite the size decrease and their counterparts in terms of pricing adjustments. The confirmation is based on individual UPC lookups and visual inspection of the package. Key metrics such as changes in size, price fluctuations, and sales variations before and after the size reduction event are examined. Standard deviations are indicated in parentheses.

D Alternative Specifications for Estimating Consumer Responses

D.1 Individual Coefficients

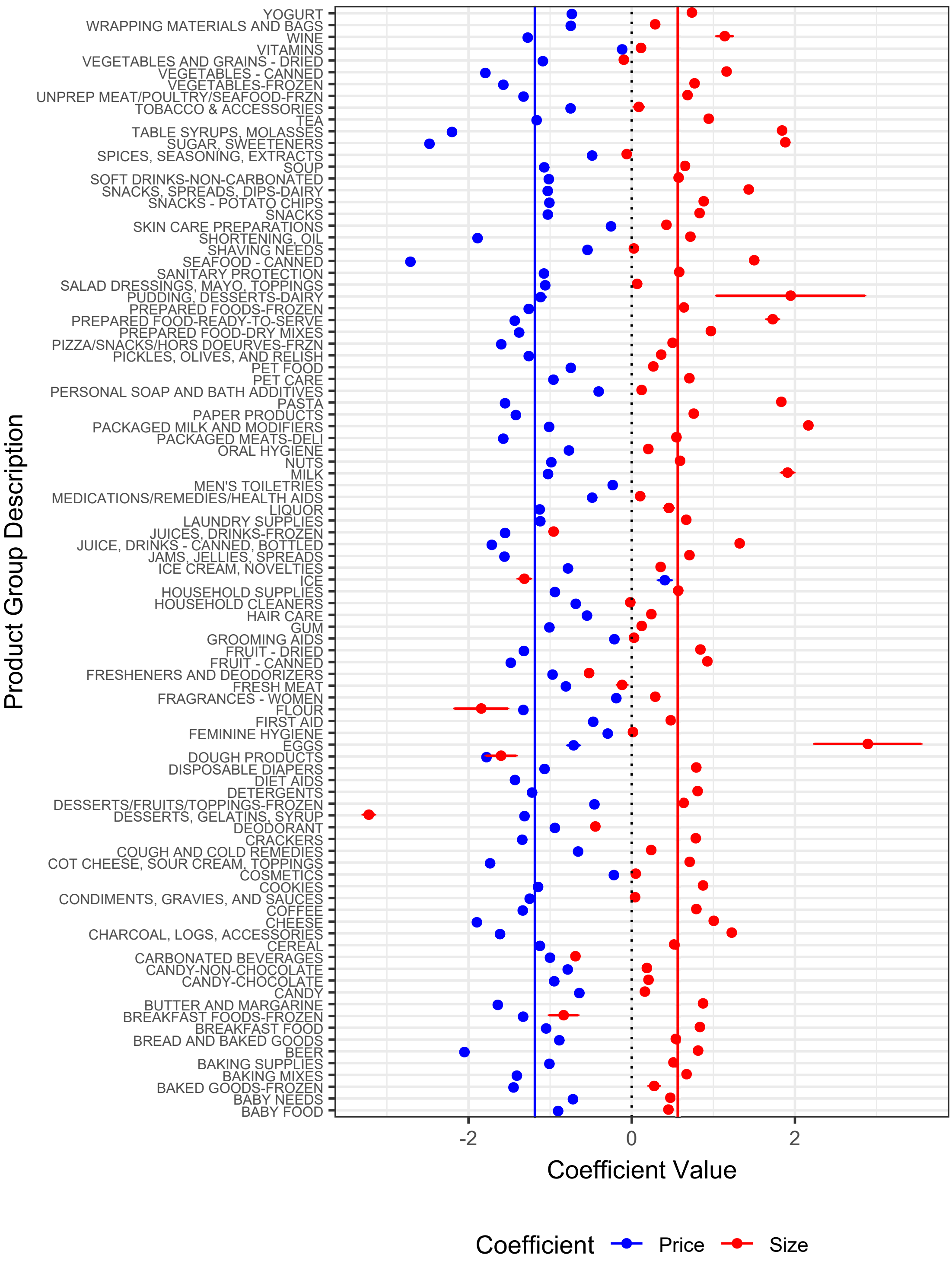
Here, we give an overview of the estimation results by product group. Figure D.1 shows the estimated price and size elasticity of the benchmark model using a linear least squares regression. The figure corresponds to Figure 7, which plots the histogram of those coefficients.

D.2 More Granular Fixed Effects

Demand across product-store combinations often fluctuates over time, driven by factors such as changes in store assortments or competitor price reactions. In our benchmark model, we use store-product fixed effects that remain constant across years. This approach can lead to biased estimates, as we aggregate over multiple years, which does not account for time-varying changes. To address these concerns, we estimate a specification incorporating store-product-year fixed effects, allowing product-store effects to vary over time. Following the baseline model of DellaVigna and Gentzkow (2019), this specification also includes product-store-week-of-year fixed effects to capture seasonal variations. The estimates from this specification, presented in Figure D.2a, remain robust and consistent with our main findings. The estimated average price elasticity is -1.24, and the average size elasticity estimate is equal to 0.62.

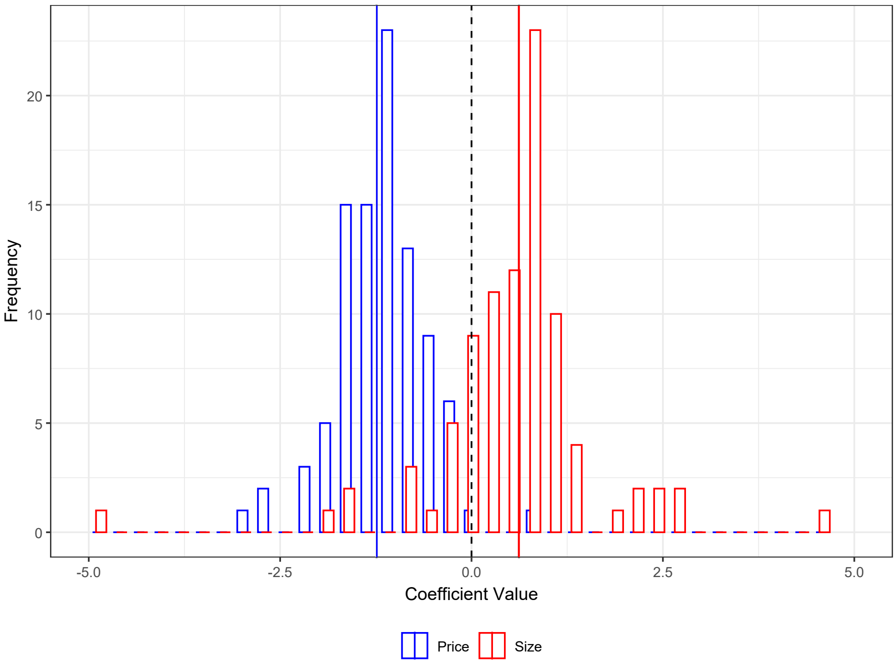
The fixed effects in our benchmark model imply that we also consider size variations from products not classified as downsized or upsized at the store level. To address concerns that these elasticities may not reflect actual downsizing or upsizing, we estimate an alternative specification with UPC-store fixed effects for products not meeting our down- or upsizing criteria and with product-store fixed effects, as in our benchmark model, for downsized and upsized products. These fixed effects account for all size variations linked to a UPC change that do not meet our definition of downsized or upsized products. The estimated elasticities are presented in Table D.2b, yielding a nearly equivalent price elasticity (-1.18) and a near-zero average size elasticity of 0.03. These findings align with our interaction term specification, which shows negligible size elasticity estimates for downsized products. Together, the robustness checks with more granular fixed effects further support our conclusion that consumers underreact to size changes relative to price variations.

Figure D.1: Estimated Unit Price and Package Size Elasticities for Different Product Groups

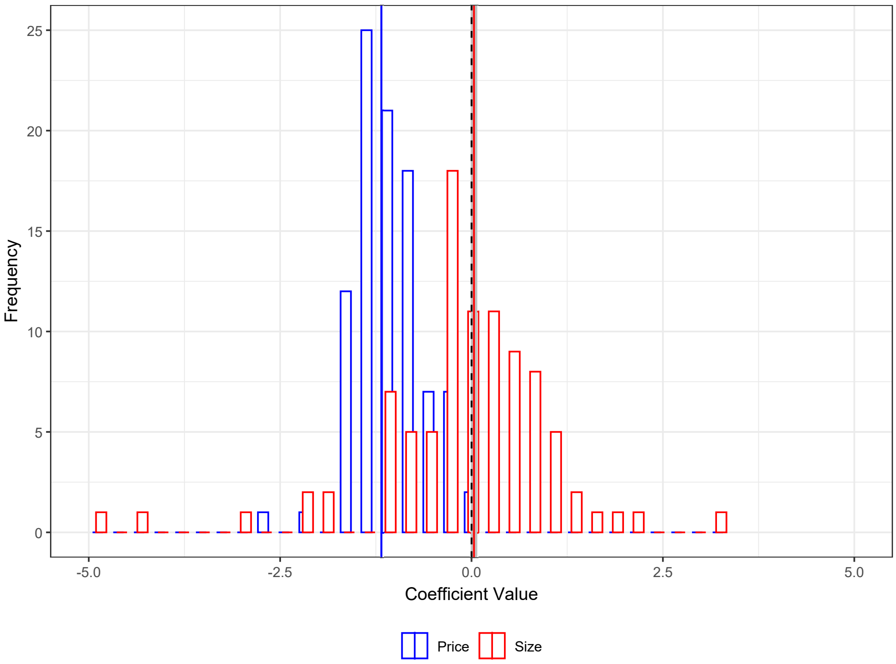


Notes: This graph presents the estimated unit price elasticities (in blue), package size elasticities (in red), and corresponding 95% confidence intervals in different product groups, according to equation 1. The coefficients, therefore, refer to $\hat{\eta}_p$ *and* $\hat{\eta}_l$ *in model 1. The solid blue line (red line) illustrates the sales-weighted average unit price (package size) elasticity across all product groups. Weights are based on the sales within product groups. Error bars denote the 95% confidence intervals of the weighted coefficients. For visibility, we exclude product groups with extremely large standard errors that result in rejecting the hypothesis that the point estimates are within the range of* ≤ -5 *or* ≥ 5*. However, those product groups are part of the weighted average calculations.*

Figure D.2: Estimated Unit Price and Product Size Elasticities with more granular fixed effects



(a) Store-product-year and seasonal fixed effects



(b) UPC-level fixed effects

Notes: The graphs present histograms of the estimated unit price elasticities (in blue) and package size elasticities (in red), according to equation 1, across product groups. In contrast to our benchmark model, we apply product-store-year and product-store-week-of-the-year fixed effects instead of week and product-store fixed effects in Subfigure D.2a. In Subfigure D.2b, we apply product-store fixed effects for downsized and upsized products and UPC-store fixed effects for all other products. The solid blue line (red line) illustrates the sales-weighted average unit price (package size) elasticity across all product groups. Weights are based on sales within product groups. The gray area around the weighted mean denotes the 95% confidence intervals of the weighted coefficients. For visibility, we exclude product groups with extremely large standard errors that result in rejecting the hypothesis that the point estimates are within the range of ≤ -5 *or* ≥ 5*. However, those product groups are part of the weighted average calculations.*

D.3 Quarterly data

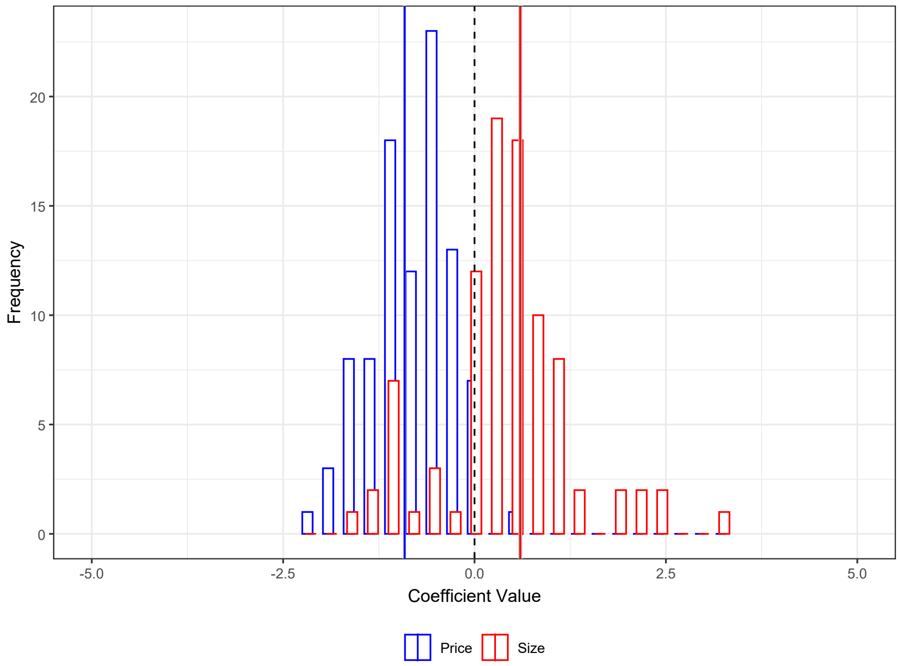
To assess whether the observed underreaction to size changes relative to price persists in the medium term, we re-estimated our benchmark model using quarterly aggregated data. For quantities, we sum all sales within a quarter at the store-product level and then take the logarithm. For prices and sizes, we calculate the arithmetic mean at the store-product-quarter level before taking the logarithm. The estimated elasticities, presented in Figure D.3, are similar to our short-term estimates. For the specification with fixed effects as in our benchmark model (Figure D.3a), the average price elasticity is –0.92 and the size elasticity is 0.59. For the alternative specification using Store-UPC fixed effects for non-downsized or upsized products (Figure D.3b), the estimated price elasticity is –0.99 and the size elasticity is close to zero (0.01). These results suggest that consumers tend to underreact to size changes even over the medium term.

D.4 Instrumental Variable Approach

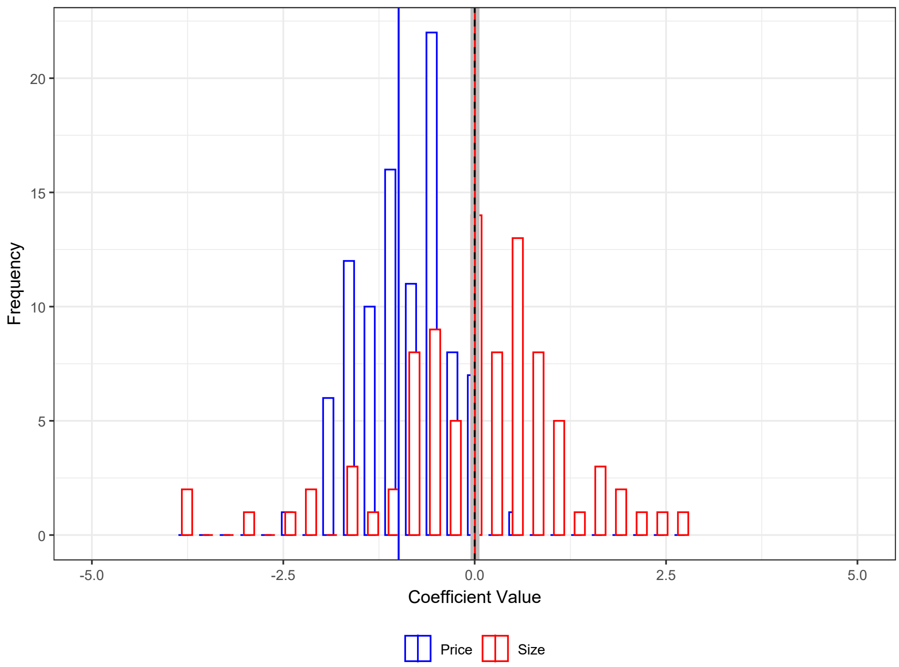
We also use an IV approach, following DellaVigna and Gentzkow (2019), where we instrument the weekly price and product size in store s with the average prices and product sizes from other stores in s's chain located outside the respective store's designated market area, i.e., Nielsen DMA region(s). The key assumption here is that the timing of chain-level size and price changes is unrelated to local demand shocks after we control for fixed effects. This approach is common in the industrial organization and marketing literature (Hausman, 1996; Nevo, 2001). However, it has faced strong and justified criticism (Rossi, 2014), and the results should be interpreted with caution.

Figure D.4 presents the results for the IV regressions. In the specification with store-product and week fixed effects, as in the benchmark model (Subfigure D.4a), the weighted average price elasticity is -1.73 and the package size elasticity is 0.82. The first stage is strong in most cases, and Section D.4.1 provides further details, including the relationship between size estimates and first-stage F-statistics. As an additional robustness check, we estimate an IV model with store-product-year and store-product-week-of-the-year fixed effects, as in DellaVigna and Gentzkow (2019). The results, shown in Subfigure D.4b, are consistent with our main estimates and the other IV specification. Overall, the IV estimates validate the conclusion that consumers underreact to price changes stemming from package size reductions than to direct unit price changes.

Figure D.3: Estimated Unit Price and Product Size Elasticities - Quarterly data



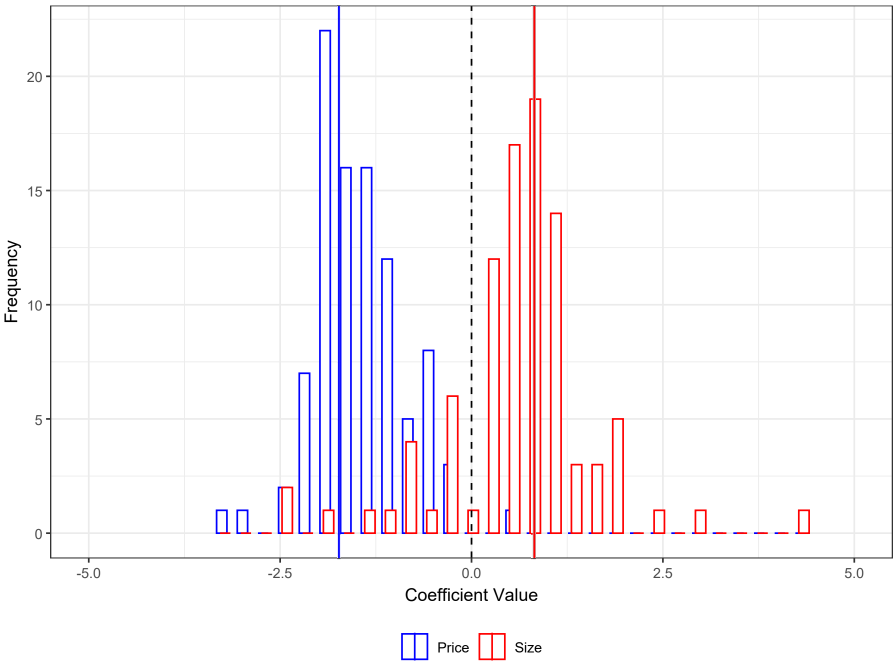
(a) Quarterly data - Benchmark



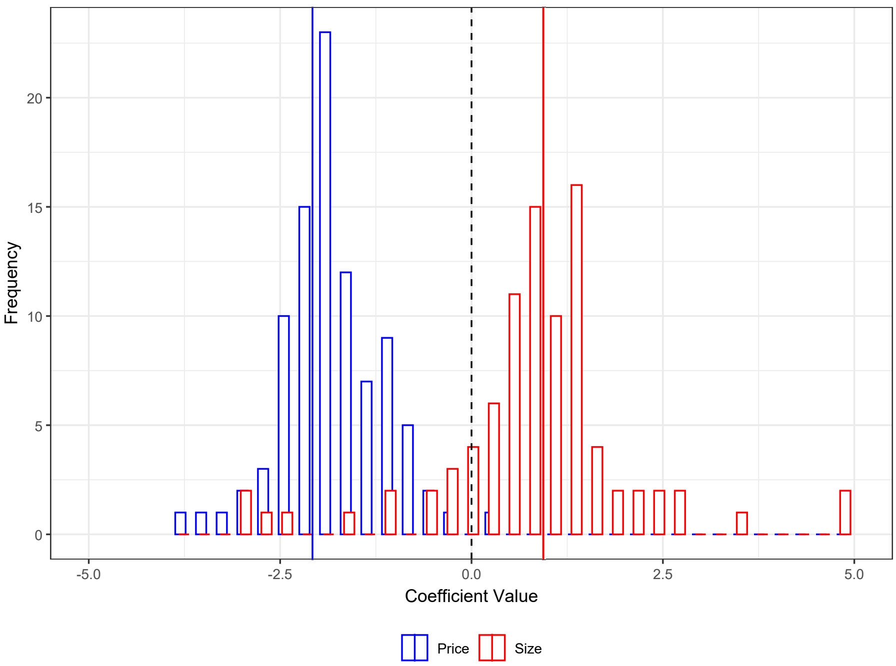
(b) Quarterly data - UPC-level fixed effects

Notes: The graphs present histograms of the estimated unit price elasticities (in blue) and package size elasticities (in red), according to equation 1, across product groups using quarterly aggregated data (Subfigure D.3a). In Subfigure D.3b, we apply product-store fixed effects for downsized and upsized products and UPC-store fixed effects for all other products. The solid blue line (red line) illustrates the sales-weighted average unit price (package size) elasticity across all product groups. Weights are based on sales within product groups. The gray area around the weighted mean denotes the 95% confidence intervals of the weighted coefficients. For visibility, we exclude product groups with extremely large standard errors that result in rejecting the hypothesis that the point estimates are within the range of ≤ -5 *or* ≥ 5*. However, those product groups are part of the weighted average calculations.*

Figure D.4: IV-Estimated Unit Price and Product Size Elasticities



(a) Main model



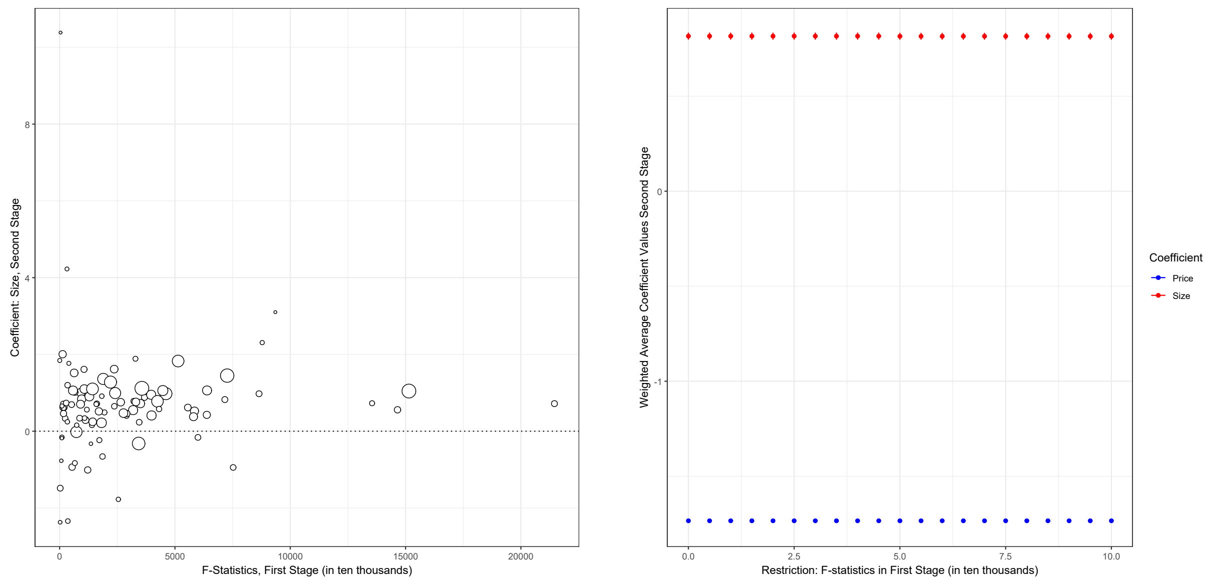
(b) Store-product-year and seasonal fixed effects

Notes: The figures present histograms of the estimated unit price elasticities (in blue) and package size elasticities (in red) across product groups, using an IV approach. We instrument weekly price and the weekly package size in store s with the average of prices and package sizes across other stores in s's chain that are located outside the respective store's Nielsen DMA region(s). Subfigure D.4a uses store-product and week fixed effects, as in the benchmark model, and Subfigure D.4b uses store-product-year and store-product-week-of-the-year fixed effects, as in DellaVigna and Gentzkow (2019). The solid blue line (red line) illustrates the sales-weighted average unit price (package size) elasticity across all product groups. Weights are based on the sales within product groups. The gray area around the weighted mean denotes the 95% confidence intervals of the weighted coefficients. For visibility, we exclude product groups with extremely large standard errors that result in rejecting the hypothesis that the point estimates are within the range of ≤ -5 or ≥ 5. However, those product groups are part of the weighted average calculations.

D.4.1 First Stage

In our IV analysis, we use two primary instruments: averages of weekly in-store prices and product sizes across other stores in the same chain but located outside the respective store's Nielsen DMA region(s). The first-stage F-statistics (i.e., from the regression of product sizes in a store on the average sizes of the same product in other stores of the same chain in other markets) are generally very large, consistent with the findings of DellaVigna and Gentzkow (2019). To assess the strength of the first-stage and its influence on our elasticity estimates, we conduct two analyses. First, Figure D.5a illustrates the relationship between the first-stage F-statistic for the size instrument and the second-stage size elasticity estimate. Each point on the graph represents a product group, with point sizes proportional to sales. While outliers are more common among product groups with relatively low F-statistics, Figure D.5b shows that the second-stage average elasticity estimates remain stable when excluding product groups with low (adjusted) F-statistics. In detail, we exclude product groups based on first-stage F-statistic thresholds, as indicated on the x-axis, and then compute the corresponding sales-weighted average estimates for size elasticities (shown in red) and price elasticities (shown in blue) for the remaining product groups.

Figure D.5: Evaluating First Stage of Instrumental Variable Regression



(a) F-statistics and Size Coefficient

(b) Different F-statistic Restrictions

Notes: Subfigure D.5a plots the relationship between first-stage F-statistics and second-stage size elasticity estimates. Each point represents a product group, with point sizes proportional to sales. For clarity, we exclude product groups with extremely high F-statistics. The excluded size elasticity estimates are small and are no outliers. Subfigure D.5b excludes product groups based on first-stage F-statistic thresholds shown on the x-axis. For the remaining product groups, we then calculate sales-weighted average size (red) and price (blue) elasticities. The error bars indicate 95% confidence intervals.

D.5 Alternative Interaction Term specifications

Figure D.6 shows differences in price and size elasticities between downsized and non-downsized products using two alternative model specifications. The first specification (Subfigures D.6a and D.6b) follows the benchmark model, incorporating store-product and week fixed effects interacted with the downsized product indicator. This specification already shows a stronger underreaction to size relative to price, even for non-downsized products (Subfigure D.6a). The difference in average size elasticity between downsized and upsized products is significant but smaller (around 0.07), as indicated by the red line in Subfigure D.6b. This smaller difference is expected, as the estimates also capture broader size variations unrelated to direct downsizing and upsizing, which dilutes the estimated difference.

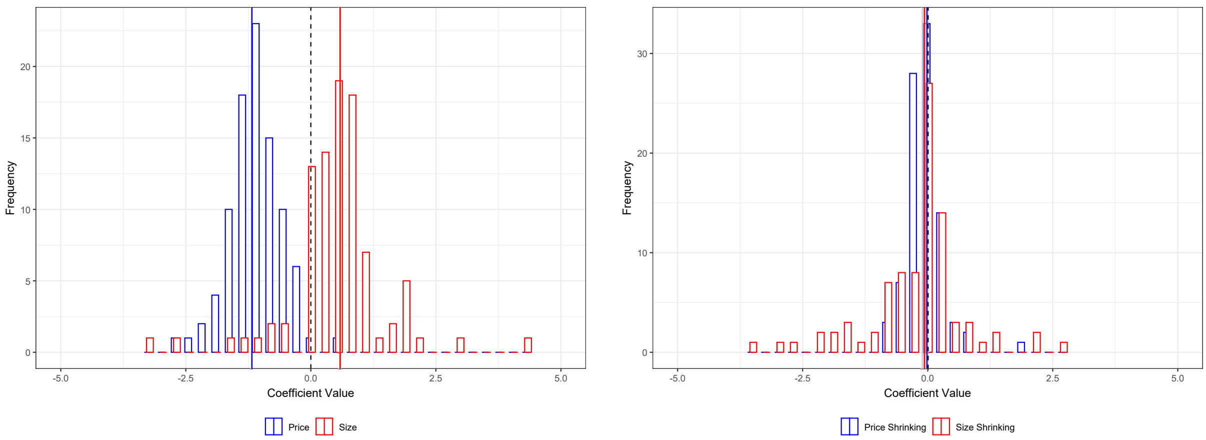
The second specification (Subfigures D.7a and D.7b) incorporates more granular store-product-year and store-product-week-of-the-year fixed effects, as in DellaVigna and Gentzkow (2019). These results show a more pronounced difference in size elasticities between downsized and upsized products. In sum, our main findings are reaffirmed: size elasticity is significantly higher for upsized products, and price elasticities remain comparable between downsized and upsized products.

E Additional Figure

Figure E.1 illustrates changes in purchase volume across product groups before and after a size change. We calculate the total change in purchase volume by comparing purchases in the 52 weeks before and after a size change at the store-product level. This difference is then expressed as a percentage of pre-change purchase volume. We average these relative changes within each module and aggregate them to the product group level using sales weights.

On average, purchase volume decreases around 3.5% for downsized products. However, this decline is not due to fewer units sold, as unit sales slightly increase, but rather to the reduction in product size. This pattern aligns with rising sales (Figure 6), which suggests limited substitution away from downsized products, at least within the first year after a size change. For products we increase in size we find an increase in purchase volume of around 20%.

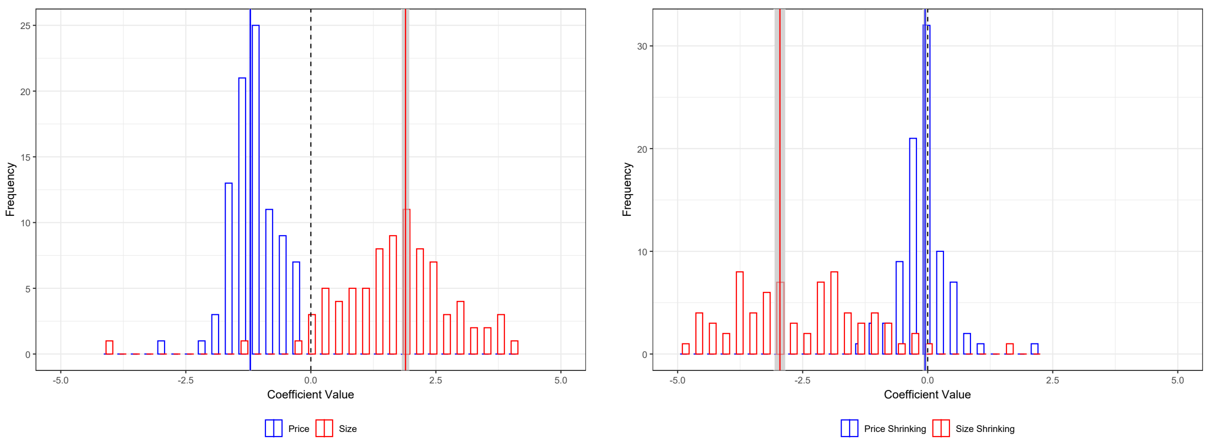
Figure D.6: Price & Size Elasticities with Interaction Term - Store-product fixed effects



(a) Price and size elasticities for upsized products

(b) Interaction terms for downsized products

Figure D.7: Estimated Unit Price and Product Size Elasticities with Interaction Term - Store-product-year and seasonal fixed effects

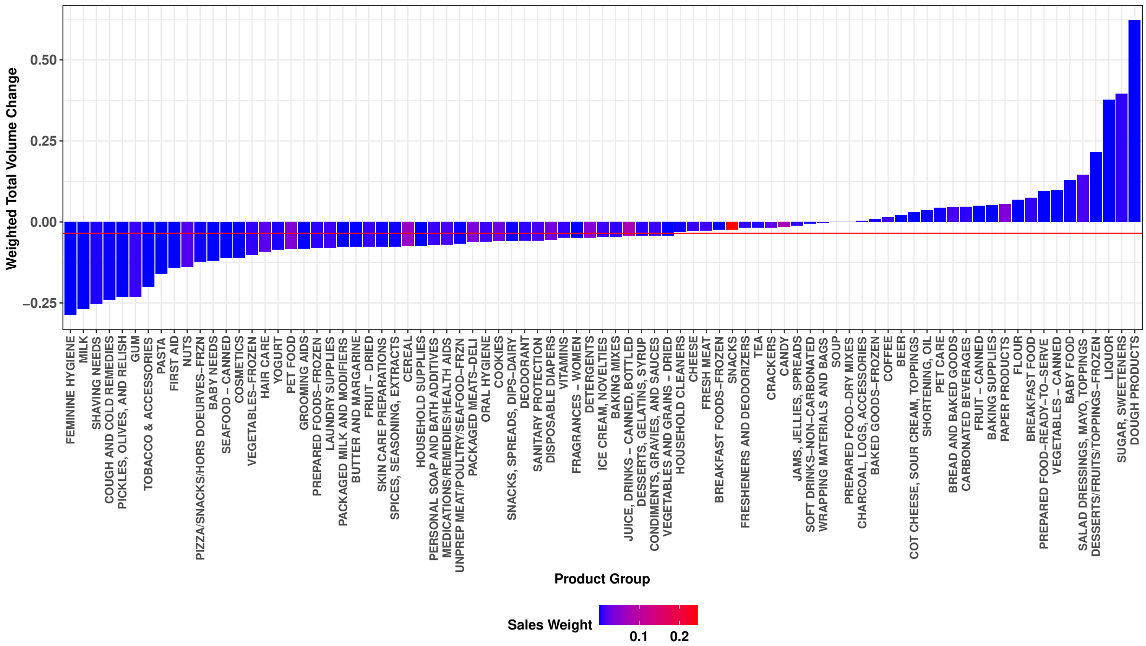


(a) Price & size elasticities - upsized products

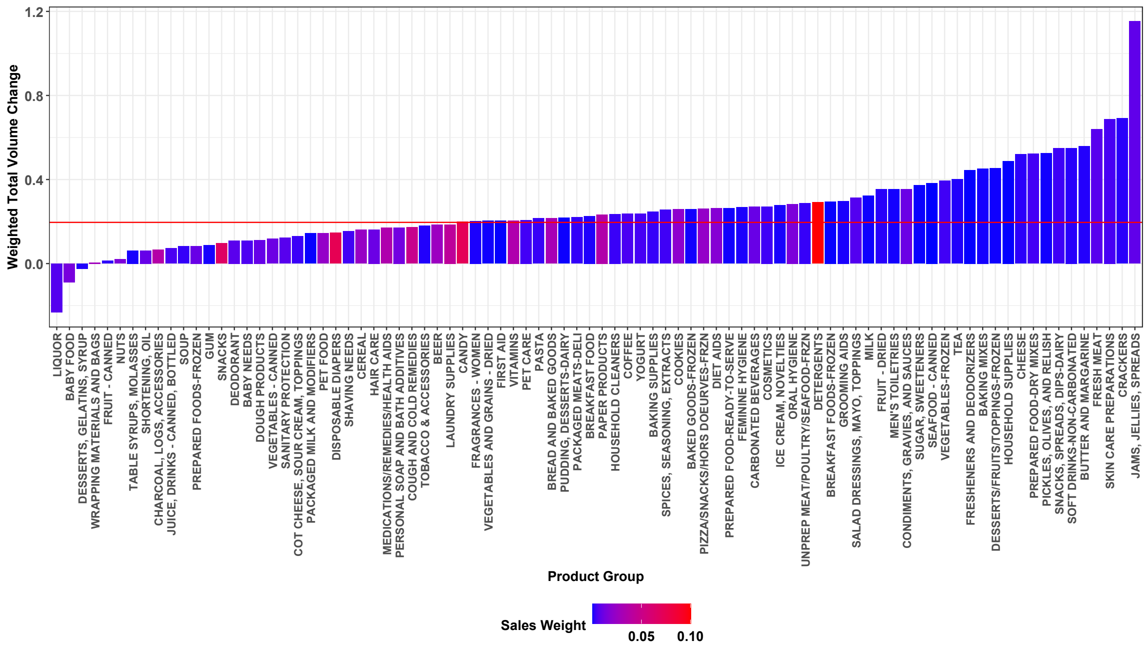
(b) Interaction terms - downsized products

Notes: Figure D.6 presents the results of two alternative specifications with an interaction term for downsized products, based on equation 2. Subfigures D.6a and D.6b display estimates from models that use store-product and week fixed effects (interacted with the downsized product indicator), similar to the benchmark model. Subfigures D.7a and D.7b show estimates from a model incorporating store-product-year and store-product-week-of-the-year fixed effects. Subfigures D.6a and D.7a present histograms of the estimated unit price elasticities ($\hat{\eta}_p$, in blue) and product size elasticities ($\hat{\eta}_l$, in red) for upsized products across product groups. Subfigures D.6b and D.7b show histograms of the interaction terms coefficients from equation 2, interpretable as differences in unit price elasticities (in blue) and package size elasticities (in red) between downsized and upsized products. In all histograms, the solid blue line (red line) represents the sales-weighted average unit price (product size) elasticity across all product groups. The gray shaded area around the weighted mean denotes the 95% confidence intervals of the weighted coefficients. For visibility, we exclude product groups with extremely large standard errors that result in rejecting the hypothesis that the point estimates are within the range of ≤ -5 or ≥ 5. However, those product groups are part of the weighted average calculations.

Figure E.1: Changes of Volume Purchased across Product Groups



(a) Purchased Volume, Product Size Decrease



(b) Purchased Volume, Product Size Increase

Notes: This graph illustrates the purchased volume changes for product groups for product size decreases and product size increases. Subfigure E.1a investigates product size reductions, and Subfigure E.1b assesses size increases. Changes in volume are measured at the store-product level and compare sales the year before and the year after the product size change. The analysis employs an aggregation approach, consolidating data from individual products into product groups, considering their weights in sales. The color intensity of each bar in the figure corresponds to the sales weight of the product group, with deeper shades of red indicating higher sales volumes within that group. The red line corresponds to the weighted average across all product groups.

References

Stefano DellaVigna and Matthew Gentzkow. Uniform pricing in us retail chains. *The Quarterly Journal of Economics*, 134(4):2011–2084, 2019.

GS1US. Gtin management guidelines, net content. https://www.gs1.org/1/gtinrules/en/rule/266/declared-net-content, 2023. Retrieved 3 January 2024.

Jerry A Hausman. Valuation of new goods under perfect and imperfect competition. In *The Economics of New Goods*, pages 207–248. University of Chicago Press, 1996.

Günter J Hitsch, Ali Hortacsu, and Xiliang Lin. Prices and promotions in US retail markets. *Quantitative Marketing and Economics*, 19(3):289–368, 2021.

Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2): 307–342, 2001.

Peter E Rossi. Even the rich can make themselves poor: A critical examination of iv methods in marketing applications. *Marketing Science*, 33(5):655–672, 2014.