Dimensions of Retail Price Competition and Consumer Choice Constraints*

Kristina Brecko[†] Paul Ellickson[‡] Avery Haviv[§]

May 7, 2025

Abstract

In this paper, we examine differentiation in grocery retail pricing strategies, focusing on their appeal to diverse consumer segments and their implications for the most constrained consumers. Our findings reveal that the grocery industry is highly fragmented, with distinct pricing strategies tailored to varying consumer needs and budget constraints, supported by concurrent investment strategies determining assortment and store location. We show that lower-income households predominantly shop at retailers offering lower and more stable prices, while higher-income households with greater budget flexibility gravitate toward retailers that provide quantity discounts and deeper intertemporal price promotions. These choices cannot be fully attributed to differences in assortment preferences or travel constraints; rather, they are shaped by binding budget constraints, as reflected in the savings consumers forgo when shopping at different retailers. Finally, we demonstrate that retailers' costly commitment to distinct pricing strategies creates healthful assortment trade-offs for the most budget constrained households and discuss implications for policy design.

*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. The authors thank their colleagues at Rochester, seminar participants at UC Berkeley and UCLA Marketing Camp, as well as conference participants at the Choice Symposium and Marketing Science conferences for the helpful comments.

[†]Simon Business School at University of Rochester, kristina.brecko@simon.rochester.edu

[‡]Simon Business School at University of Rochester, paul.ellickson@simon.rochester.edu

[§]Simon Business School at University of Rochester, avery.haviv@simon.rochester.edu

1 Introduction

Past research has shown that product prices and promotions are typically uniform within a retailer chain – largely due to the complexity of distinguishing differences in demand and managing item-level pricing – but vary substantially across chains (DellaVigna and Gentzkow (2019), Hitsch et al. (2021), Chintagunta et al. (2003)). These cross-retailer price differences create a complexity for consumers who must form an evaluation of a retailer's price offering in choosing where to shop. A retailer's expensiveness depends on a consumer's shopping needs, as retailers vary in their relative expensiveness across different products (Clerides et al. (2023)). Additionally, it is shaped by the availability and depth of intertemporal and volume discounts, as well as the household's ability to take advantage of these savings to reduce overall costs. Understanding how consumers evaluate this pricing complexity is key to understanding their store choice decisions.

In this paper, we examine how consumer price evaluations depend not only on their needs, preferred products, and the retailer's price levels but also on price variation, available discounts, and *household constraints* that may limit their ability to benefit from these savings. We demonstrate that retailers in the grocery market have adapted their *pricing strategies* to target differences in consumer preferences, price sensitivities, *and constraints*, driving the fragmentation and diversification of pricing approaches that have emerged in this market. We further show that retailers commit to pricing strategies through costly investments in assortment and location, reinforcing differentiation and enabling them to sustain distinct pricing strategies, even in a market where the same products are often available across retailers. This targeted differentiation creates trade-offs for the most constrained households, potentially contributing to nutrition gaps between the lowest- and highest-income consumers. It also generates externalities – akin to the preference externalities first documented by Waldfogel (2003) – for households whose constraints differ from those of their neighbors, with additional implications for the most constrained consumers.

We first develop four continuous pricing strategy measures that collectively capture the pricing strategies used by dominant U.S. grocery retailers: average expensiveness (*relative price*), lowest monthly price (*minimum monthly price*), minimum available unit price (*minimum unit price*), and minimum available volume price (*minimum volume price*). These focal pricing strategies are motivated by their varying appeal to consumers with different constraints, including budget limitations, storage capacity, and travel and time constraints. These factors have been shown to affect consumers' ability to take advantage of discounts in certain product categories (Orhun and Palazzolo (2019)) and have been theorized as targets of specific pricing strategies in the grocery industry (Lal and Rao (1997)).

Prior research on retailer pricing strategies has primarily examined a subset of pricing strategies used by certain retail formats – such as everyday low pricing (EDLP) and promotional pricing (HiLo) in grocery stores (Lal and Rao (1997), Bell and Lattin (1998), Ellickson and Misra (2008)) – and have largely relied on retailer self-reports.[1] Our approach systematically characterizes retailer pricing strategies using granular price data, capturing variation across all major retailers and retailer formats. Using NielsenIQ Retail Scanner data and Consumer Panel Data from the Chicago Booth Kilts Center for Marketing, we estimate the four pricing strategy measures for each retailer-category-year, covering all major U.S. retailers and approximately 900 product categories from 2012 to 2020.

Our estimates provide several key descriptive insights. First, they offer the first systematic evidence of differences in pricing strategies across retail formats. Notably, while dollar stores maintain low minimum unit prices, they exhibit higher relative prices on overlapping assortments and higher volume prices, confirming reports that shopping at dollar stores entails a real monetary trade-off (Rogelberg (2024)). Warehouse clubs offer low minimum volume prices but require significant expenditure per trip and tend to price overlapping products higher than other retailers. Discount stores as a group are the most relatively inexpensive on overlapping assortments but provide few monthly discounts and have higher volume prices than warehouse clubs, while also requiring significant per trip spending. Drug stores exemplify HiLo pricing, yet even their lowest monthly prices remain higher than all but warehouse clubs.

Importantly, we find that grocery stores span nearly the full range of pricing strategies of other retailer formats, with some resembling dollar stores and others aligning more closely with discount stores. More broadly, we document substantial variation in pricing strategies both across and within retail formats, highlighting the importance of accounting for differences be-

[1]In contrast, research documenting within-retailer uniformity in price levels and price promotions (DellaVigna and Gentzkow (2019), Hitsch et al. (2021)) connects more directly to prices paid but focuses on individual product pricing rather than the broader pricing strategies retailers employ.

tween retailers. Additionally, we find substantial variation across product departments within the same retailer, underscoring that households are likely to evaluate retailer pricing strategies differently based on their shopping needs. Consistent with prior research, we find little geographic variation in retailer pricing strategies.

We next show that differences in retailer pricing strategies are reinforced by complementary assortment and location differentiation, further amplifying their targeted nature. Lal and Rao (1997) propose that HiLo and EDLP pricing cater to distinct consumer segments and require concurrent investments in service quality. Focusing on price levels, Ellickson (2006) models grocery retailer differentiation as an outcome of endogenous quality investments, where market structure gives rise to high-quality, high-price chains and a fringe of lower-quality, lower-price independent rivals, each serving distinct consumer segments.

Our findings extend these insights across the full range of pricing strategies and retailer formats, revealing strong links between pricing strategies and assortment choices that support costly, targeted differentiation. Retailers carrying high-cost specialty departments tend to have higher prices. Fresh produce plays a key role: retailers offering random-weight fresh produce are more expensive overall, while those carrying barcoded fresh produce require higher per trip spending through higher minimum prices but provide more stable pricing. Retailers with lower and more stable prices tend to offer a broader category assortment but achieve these prices by limiting higher quality name-brand selection. Lastly, retailers with larger package sizes have significantly lower relative and volume prices with more pricing stability, whereas those with smaller package sizes maintain price stability and lower unit prices.

Retailers' joint pricing and location decisions further reinforce these patterns. We find that lower and *more stable* prices are more prevalent in lower-income areas and regions with other indicators of heightened household constraints. Additionally, we show that retailer co-location is highly localized, with ZIP-level characteristics more strongly associated with retailer pricing strategies than county-level characteristics. These findings align with previous research on the geographic targeting of EDLP and HiLo pricing strategies (Ellickson and Misra (2008)). Moreover, the findings suggest that particular pricing strategies require sufficient scale to be viable, which may lead retailers to jointly choose pricing strategies and location to ensure enough demand to support their targeted offerings.

Next, we demonstrate that even after accounting for other non-price aspects of retailer differentiation, consumers' retailer visits remain highly responsive to their *pricing strategies*, suggesting that household constraints play a significant role in retailer choice. We compute household-specific evaluations of pricing strategies and assortments based on their shopping needs. Then using a regression specification that controls for assortment evaluations and retailer accessibility within the ZIP code, we find that the most constrained households – lowest income and largest size – disproportionately visit retailers with lower and more stable pricing, a finding that aligns with previous research showing that larger basket shoppers prefer EDLP pricing (Bell and Lattin (1998)). Additionally, we find that the most constrained households disproportionately visit retailers with lower minimum prices, such as dollar stores and similar retailers. This pattern underscores the costliness of household constraints as these retailers typically offer lower minimum unit prices but higher relative and volume prices.

Our analysis of monetary savings foregone within retailer suggest that these differences in retailer visit patterns are driven primarily by binding budget constraints. A within-retailer and ZIP code comparison of foregone basket savings across households reveals that the lowest-income and largest households are the most limited in their ability to utilize inter-temporal and volume discounts, with each type of forgone savings amounting to approximately 1% of annual income and 6-7% of annual at-home food budget (Martin (2024)) for the lowest-income group. Less constrained households largely forgo fewer intertemporal and volume savings both in absolute terms and as a share of income. This finding runs counter to a prediction that more price-sensitive lower-income households would capture greater savings, and instead suggests that stable, low unit prices disproportionately appeal to these shoppers due to budget constraints that bind at the trip level. In contrast, the lowest-income and largest households do not forgo more brand and retailer savings than higher-income households. These findings suggest that travel and time constraints as well as willingness to trade-off brand name are not the primary barrier preventing the most constrained households from achieving basket savings.

Finally, we examine the implications of retailer differentiation for the most constrained households. The fragmentation in this market – designed to cater to differences in preferences, price sensitivities and constraints – creates monetary trade-offs and reduced access to healthful assortments for the most constrained households. Specifically, retailers in the bottom quintile of minimum available prices, which are disproportionately frequented by these households, tend to have a higher distribution of relative prices, with median and mean relative expensiveness at the 29^{th} and 35^{th} percentiles, respectively. These retailers also provide reduced fresh food access along the extensive and intensive margins. Only 63% of the lowest-unit-price retailers carry random-weight fresh produce, compared to 81% among retailers in higher price tiers. Additionally, the median retailer in the bottom quintile offers zero barcoded fresh produce categories, whereas the median retailer in the top quintile carries 20. These findings highlight the impact of the joint pricing and assortment differentiation strategies chosen by these firms, whereby lower-priced retailers sustain low minimum prices at the expense of a diverse produce selection, which is costly to carry. As a result, even if the most constrained households prefer healthful foods, their budget constraints may lead them to shop at stores with fewer and lower-quality options, ultimately reducing their likelihood of purchasing these products.

Furthermore, the localized nature of the retail offerings may lead to externalities – similar to preference externalities (Waldfogel (2003), Handbury (2021)) – for those residing in areas with different constraints than their own, as the pricing strategies of local retailers cater to the locally dominant consumer group. For the most constrained households residing in areas with less constrained households, this entails a trade-off between adhering to a constrained budget and access to grocers with a healthful product offering.

We conclude by discussing the implications of our findings for the ongoing debate about food desserts – areas without direct access to grocery stores. Our research rationalizes food desserts as an equilibrium outcome, resulting from fragmentation in retailer pricing strategies in response to differences in consumer constraints. While prior work has highlighted the role of demand in shaping grocery access, it has largely attributed these patterns to differences in preferences for healthful food, without accounting for household constraints (Allcott et al. (2019)). Our findings suggest that even when preferences for healthful foods are similar, households with tighter constraints may trade-off healthful assortments in order to better satisfy overall basket needs within a limited budget. This highlights that nutrition education alone is likely insufficient to close nutritional gaps, unless paired with basket-level subsidies at retailers offering healthful assortments. Additionally, supply-side policies aiming to improve grocery access in food deserts should take into account both the presence of fresh produce in retailers' assortments and their pricing strategies to ensure meaningful improvements in access for the most constrained households.

As described above, we view our study as contributing primarily to the literature on retail pricing strategies, their targeted nature and the concurrent investments required to sustain them (Lal and Rao (1997), Bell and Lattin (1998), Ellickson (2006), Ellickson and Misra (2008)) as well as the broader literature on retail pricing behavior and its impact on consumers, including zone and uniform pricing (Chintagunta et al. (2003), Adams and Williams (2019), DellaVigna and Gentzkow (2019), Hitsch et al. (2021)), price dispersion (Kaplan and Menzio (2015), Kaplan et al. (2019), Clerides et al. (2023)) and the cyclic nature of price promotions (Bronnenberg et al. (2006)). To this literature, we contribute a systematic and data-driven study of the rich literature on retailer *pricing strategies*, the differentiation that underpins it and the implications this differentiation has for the most constrained consumers.

In considering how nuances in consumer price evaluation shape retailer pricing strategies, we also relate to Clerides et al. (2023), who consider store expensiveness as a function of the consumer's basket and Thomassen et al. (2017), who develop a multi-category model of consumer store choice to examine how retailer pricing responds to complementarities across product categories in consumer baskets. We contribute by showing that consumer retailer choices are shaped by consumer constraints and retailers in the U.S. grocery market have adapted their pricing strategies to target constraints along with preferences and price sensitivities.

We contribute to the literature on the sources of nutritional disparities across income groups (Allcott et al. (2019), Caoui et al. (2022)), the distributional consequences of retail closures (Cao et al. (2024)), and the role of non-homothetic demand in shaping cost-of-living differences across income levels (Handbury (2021)). While prior work has largely attributed differences in grocery shopping patterns and nutritional intake to consumer preferences, we provide a more nuanced explanation by highlighting the role of household constraints. Specifically, we propose that differences in food choices and retailer selection across income groups can, in part, be rationalized by how static household budget constraints impact consumer choice of retailer.

Finally, our findings may help explain the persistent differences in grocery retail landscapes between developed and developing economies. While large chains of big-box retailers dominate developed markets, their penetration and growth have lagged in developing markets, while small independent stores persist (Child et al. (2015)). This divergence may stem from a mismatch between the pricing strategies and assortments offered and the needs of heavily constrained consumers who constitute a larger proportion of these markets.

The rest of the paper proceeds as follows. Section 2 introduces the data used in this study. Section 3 develops, estimates, and analyzes the variation in four pricing strategy measures that characterize the U.S. grocery retail landscape and potentially cater to households with differing budget constraints. Section 4 demonstrates that these different pricing strategies involve complementary investments in assortment and location to sustain their viability. Section 5 shows that consumers select stores in ways consistent with their budget constraints, even after accounting for heterogeneous preferences over assortments and travel costs. Section 6 discusses how this costly differentiation and targeting of consumer budget constraints affects the welfare of the lowest-income households, and Section 7 concludes.

2 Data

We use the NielsenIQ Retail Scanner Data (RMS) and NielsenIQ Consumer Panel Data (HMS) provided by the Chicago Booth Kilts Center for Marketing in our analysis. We focus our analysis on the 2012-2020 data sets, as the retailer codes, which are integral to our analysis as described below, are less frequently populated in the RMS data in years prior to 2012. We supplement these data with demographics data from the U.S. Census Bureau American Community Survey.

2.1 Data Description

The RMS data set records store-level weekly quantity sales and average prices paid for each universal product code (UPC) at participating retailers. In addition to the sales data, the RMS data set provides supplementary information on store characteristics, such as the store location (ZIP3 and county), store format (e.g., grocery, discount, etc.) and an anonymized identifier for the retail chain to which the store belongs. The HMS data set tracks details about participating panelist households' shopping trips and purchases, including the identity of the store or retailer visited,[2] the date of the shopping trip, as well as the quantity and price of each UPC purchased on the shopping trip. These data also contain information about panelist and retailer characteristics, including household income and county of residence, as well as the retailer format. The key difference between the RMS and HMS data sets is that the former gives a relatively complete insight into weekly UPC sales and prices at each participating store, whereas the latter gives a relatively complete insight into the UPC purchases of a given household, but not any given store.

We use these data for two purposes: (1) to construct price indices and (2) to analyze household choice of where to shop and which products to buy. The rich price panel of the RMS data set serves as our primary data source in constructing the price indices of participating retailers. Because a number of large retail chains do not participate in the RMS data ("non-participant retailers"), the RMS data are not sufficient to characterize the price indices of all the retailers in a given household's choice set. Consequently, for our retailer visits analyses, we construct price indices for non-participant retailers by supplementing the RMS data with the UPC prices observed to be paid at these retailers in the HMS data.

There are two key drawbacks to using the HMS data in this manner. The first is that the HMS data offer a much more sparse view into the weekly prices of a given UPC at a given retailer. Even if the observed prices were effectively sampled randomly, using these prices to infer the unobserved prices and consequently price indices is likely to lead to additional measurement error relative to the RMS data. Moreover, because the price observations in the HMS data are likely to be systematically selected, price indices based on these data may be skewed relative to the RMS-based measures. The first concern introduces classical measurement error, leading to attenuation bias in our estimates of the relationship between store pricing strategies and household store choice and shopping behavior. To address the second concern, we (1) test the robustness of our results by limiting the analysis to RMS-participating stores only (Section 5), and (2) compare price indices for retailers present in both RMS and HMS data, estimated using (a) RMS data alone and (b) the combined RMS and HMS data, as applied in our main approach.

The detailed HMS household shopping trip and purchase data also serve as the basis for our

[2] In the HMS data, only stores of RMS-participating retailers have a unique store identifier. All other stores are identified by the anonymized retailer identifier only.

analysis of households' choices to shop at retailers with particular pricing strategies and their choices of products at these stores. We additionally form two measures of household income using the mid-point of the household income range provided by NielsenIQ. We assign each household to a year-specific household income quintile by determining whether the midpoint of the household income range in that year falls into a given year's household income quintile, as defined by Tax Policy Center (2024).[3]. We also compute a continuous household income measure by deflating the mid-point of the household income bucket in a given year to the January 2010 levels using U.S. Bureau of Labor Statistics (2024). Analyzing how household choices vary with income and other constraint-related demographics allows us to examine how, and which, constraints influence retailer choice and grocery spending.

2.2 Data Selection: Retailers and Products

We restrict our analysis to the five retail formats in the NielsenIQ data where grocery products constitute a substantial share of sales: Discount Store, Dollar Store, Drug Store, Grocery, and Warehouse Club.[4] We also focus on products in the most common grocery categories. The Kilts Center categorizes products hierarchically: each product belongs to a unique *product module* (e.g., refrigerated milk), which in turn belongs to a less granular unique *product group* (e.g., milk), and *product department*, the most aggregated classification (e.g., dairy). To construct the four pricing strategy measures, we use data from products in all categories within 9 grocery departments: Alcoholic Beverages, Dairy, Deli, Dry Grocery, Fresh Produce, Frozen Foods, Health & Beauty Care, Non-Food Grocery, and Packaged Meat. We exclude General Merchandise, Magnet Data (consisting of random-weight items found only in the HMS data), and products lacking department descriptions.

When calculating the price indices, we limit attention to UPCs purchased at least 1,000 times and sold in at least 200 stores, and to store-UPC observations with recorded prices in at least 39 weeks (75% of the year). The latter threshold reduces reliance on imputed prices, whereas the former is crucial for constructing relative price measures, which require overlapping

[3]Although we use year-specific income quintiles in our analyses, we display the approximate 2018 income quintiles for readability of our results tables.

[4]This excludes convenience stores in the RMS data and specialty stores (e.g., bakeries) in the HMS data—a limited concern, as households across income levels primarily shop for groceries at grocery stores and other formats within sample (Allcott et al. (2019)).

Table 1: Price Indices Data Summary

Data	Price Indices				Unique Products
	Retailer-County-Categories (1)	Retailers (2)	Counties (3)	Categories (4)	UPCs (5)
RMS Only	8,993,114	183	2,691	1,116	298,763
RMS & HMS	5,416,214	469	2,970	930	349,299

This table provides an overview of the data used to construct the price indices, focusing on the number of unique UPCs on which our price index calculations are based in column (5), and retail outlets, counties and categories for which we recover price indices in columns (2)-(4). All counts are unique counts across all in-sample years (2012-2020). Tables B.1 and B.2 complement this table by providing more detailed data summaries by year for the price indices based on RMS data only.

UPC assortments across stores.

Table 1 shows that, depending on the data set, our retailer-county-category price indices are constructed using approximately 299-349 thousand unique UPCs in 1,116-930 product categories at 183-469 unique retailers. The RMS data contain more product categories, leading to a greater number of retailer-category and retailer-county-category combinations, but cover fewer unique retailers and counties than the combined RMS and HMS dataset. In contrast, the RMS and HMS data capture a more representative set of retailers – about 2.5 times more than RMS alone – but yield fewer retailer-county-category observations due to sparser purchase coverage per retailer. Tables B.1 and B.2 summarize the RMS-only price indices by year and confirm that the covered categories account for the vast majority of total sales revenue.

3 Price Indices as Measures of Store Pricing Strategies

In choosing where to shop, households must assess the prices of the many products that make up their grocery basket. Moreover, households at different income levels face varying costs and constraints that affect their grocery shopping: budget limits, storage capacity, and travel and time costs affect households' ability to take advantage of temporary discounts or bulk purchases (Bell and Lattin (1998), Orhun and Palazzolo (2019)).[5] Thus, these differently constrained households may rely on distinct "shopping rules" when choosing where to shop, creating opportunities for retailers to target specific segments through their pricing strategies.

For example, EDLP stores with stable, predictable prices may attract households focused

[5]Web Appendix A provides a description of how these costs and constraints may vary with household income.

on minimizing expected grocery costs, but only able to purchase standard package sizes to realize those savings. HiLo retailers with frequent promotions may appeal to less constrained households who benefit from buying in bulk and whose shopping strategy involves timing purchases to stock up when relevant items are discounted. Club stores that emphasize volume discounts may appeal to households with sufficient storage and transportation capacity, seeking the best quantity discounts and unconstrained by the lumpy spending required to obtain them.[6] Dollar stores, which focus on small sizes at low prices, likely target households facing the strictest budget, storage, and transportation constraints – those that prioritize minimizing shopping *basket* costs when choosing where to shop.

In this section, we develop four continuous measures of store pricing strategies (hereafter "price indices") that allow us to capture such distinct retail value propositions geared towards the common shopping rules used by differently constrained households. Recent literature on grocery retail pricing has typically focused on individual dimensions, such as average prices (DellaVigna and Gentzkow (2019), Hitsch et al. (2021)) and price promotions (Bell and Lattin (1998), Ellickson and Misra (2008), Hitsch et al. (2021)). The four measures we develop capture the individual dimensions of pricing strategy considered in past literature and, importantly, parsimoniously characterize the main pricing strategies observed in U.S. grocery retail.

We first formally define the four price indices, providing additional computational details in Web Appendix C. We then present the resulting price indices, highlighting the main dimensions of variation in these price indices that motivate our analyses in Section 5 and onwards.

3.1 Definition

We characterize store-category pricing strategies along four dimensions: $\mathbf{P} = \{\mathbf{p}^{\text{rel}}, \mathbf{p}^{\text{mon}}, \mathbf{p}^{\text{unit}}, \mathbf{p}^{\text{vol}}\}$, where $\mathbf{P}$ is comprised of *relative price* ($\mathbf{p}^{\text{rel}}$), *minimum monthly price* ($\mathbf{p}^{\text{mon}}$), minimum unit price ($\mathbf{p}^{\text{unit}}$), and minimum volume price ($\mathbf{p}^{\text{vol}}$). We first compute granular price indices – at the store- and retailer-product category-year level, which allows us to then evaluate the variation in each of the measures. In subsequent sections, we use these granular measures to form aggregated retailer- and household-level measures, as necessary for further analyses.

[6] Our discussion focuses on the lumpy spending driven by large package sizes, though the annual membership fee, common to this retailer format, adds another hurdle likely to deter constrained consumers.

Relative Price ($\mathbf{p}_{csy}^{\mathbf{rel}}$) The first price index is based on the relative expensiveness of a given product (UPC) j in store s compared to other stores in which the same UPC is sold. Taking p_{jst} as the price of product j in store s in week t, we estimate a relative price vector $\hat{\mathbf{p}}_{csy}^{\text{rel}}$ as category-store-year fixed effects, using the following regression:

$$\log(p_{jst}) = \mathbf{p}_{csy}^{\text{rel}} + \xi_j + \varepsilon_{jst} \tag{1}$$

where c is the category to which product j belongs, y is the year of the calendar week t and ξ_j controls for product-specific price levels, thereby yielding a relative price index net of product assortment.[7]

The key estimates of interest is the vector of fixed effects $\hat{\mathbf{p}}_{csy}^{\text{rel}}$, which captures a given store's expensiveness relative to its competitors for overlapping products in each category-year. In other words, $\hat{\mathbf{p}}_{csy}^{\text{rel}}$ will be low in stores that set lower prices for the *same* products in a given category than other stores. Walmart, which has historically positioned themselves as an EDLP store (Ellickson and Misra (2008)), would likely have a low *relative price*, while higher-end grocery stores such as Whole Foods Market would likely have a high *relative price* (Traynor (2024)).

Minimum Monthly Price ($\mathbf{p}_{csy}^{\mathbf{mon}}$) The second price index helps characterize the intertemporal price variation in a store-category in the average month by looking at the lowest price offered for product j in store s and calendar month m compared to other stores in which the same UPC is sold. We estimate the minimum monthly price vector $\hat{\mathbf{p}}_{csy}^{\text{mon}}$ as a set of category-store-year fixed effects, using the following regression:

$$\min_{jsm}\big(\log(p_{jst})\big) = \mathbf{p}_{csy}^{\text{mon}} + \xi_j + \varepsilon_{jst} \tag{2}$$

where $\min_{jsm}\big(\log(p_{jst})\big)$ represents the minimum price of product j in store s and calendar month m to which week t belongs. As with the relative price index $\hat{\mathbf{p}}_{csy}^{\text{rel}}$, in computing the minimum monthly price $\hat{\mathbf{p}}_{csy}^{\text{mon}}$ we control for product-specific minimum price levels, thereby

[7] By including product fixed effects, we ensure that the price index of a store is analyzed independently of its product assortment. Products (UPCs) available at only one store do not contribute to comparisons between stores. Instead, our regression estimates price differences by comparing the same product sold at multiple stores and retailers. This method focuses on how pricing strategies differ across stores, without being influenced by differences in product selection.

ensuring that this index does not reflect differences in product assortment. In other words, $\hat{\mathbf{p}}_{csy}^{\text{mon}}$ will be low in stores that offer deeper discounts in a particular category than other stores selling the same products. As an example, Safeway and A&P are grocery retail chains that have historically self-reported to engage in predominantly HiLo pricing, and therefore would be more likely to have lower *minimum monthly price* (Ellickson and Misra (2008)).

Given the similarity in the approach, the minimum monthly price is significantly correlated with the relative price. To isolate the unique variation in minimum monthly price, we compute our final measure as the residuals from regressing minimum monthly price on relative price, addressing this correlation.

Minimum Unit Price ($\mathbf{p}_{csy}^{\text{unit}}$) The third price index is based on the absolute minimum unit price required to shop in category c in store s in year y. Similarly to the other price indices, we estimate the minimum unit price vector $\hat{\mathbf{p}}_{csy}^{\text{unit}}$ as a set of category-store-year fixed effects, using the following regression:

$$\min_{cst}\big(\log(p_{jst})\big) = \mathbf{p}_{csy}^{\text{unit}} + \varepsilon_{jst} \tag{3}$$

where $\min_{cst}\big(\log(p_{jst})\big)$ represents the minimum product price in category c, store s and week t. We do not include product controls in computing the minimum unit price, allowing assortment differences to influence this price index. In other words, the minimum unit price index $\hat{\mathbf{p}}_{csy}^{\text{unit}}$ will be lower in stores that offer cheap brands or small package sizes in a particular category. Retail chains such as Aldi, which carry few name brand products (Kelly (2023)), and dollar stores such as Dollar General, which typically carry smaller package sizes (Rogelberg (2024)), would likely have a low *minimum unit price*.

Minimum Volume Price ($\mathbf{p}_{csy}^{\text{vol}}$) The fourth price index is based on the expensiveness of a product in terms of volume price. This index accounts for differential prevalence of quantity or volume discounts offered in particular store-categories via bulk package sizes. As with the other three indices, we estimate the minimum volume price vector $\hat{\mathbf{p}}_{csy}^{\text{vol}}$ as a set of category-store-year fixed effects, using the following regression:

$$\min_{bst}\left(\log\left(\frac{p_{jst}}{v_j}\right)\right) = \mathbf{p}_{csy}^{\text{vol}} + \xi_b + \varepsilon_{bst} \tag{4}$$

where b is a product j's brand[8], v_j is a product j's package size, and hence $\min_{bst}\left(\log\left(\frac{p_{jst}}{v_j}\right)\right)$ represents the minimum *volume* price at which a given brand b is available in store s and week t. We control for brand fixed effects ξ_b to account for price level differences across brands. In practice, warehouse clubs and other stores that offer at least some products in larger package sizes would likely have a low *minimum volume price*.

Example Pricing Metrics As an example, in Table 2, consider four hypothetical stores labeled EDLP, HiLo, Dollar, and Warehouse selling two laundry detergent brands – Purex, which is generally lower priced (Palermo (2024)), and Seventh Generation, which is generally higher priced (Riss (2024)). All prices in this example are fabricated for demonstrative purposes.

The EDLP store carries both brands in 84 fl oz size, and prices them at \$10 and \$15 respectively. The HiLo store only carries 84 fl oz of Purex, and discounts it to 9 dollars for one week each month, for an average price of \$11.25. The dollar store carries both 40 and 84 fl oz Purex products at \$7 dollars and \$11 dollars respectively, while the Warehouse carries a larger 135 fl oz Seventh Generation bottle at \$18, and the 84 fl oz Purex at \$11.

In this example, the EDLP store has the lowest *relative price* because it offers the lowest expected price for the 84 fl oz Purex product, which is available at all four retailers. The HiLo store, however, has the lowest *minimum monthly price* because it offers the deepest monthly promotion on the same product, benefiting consumers with the flexibility to stock up. The Dollar store offers the lowest *minimum unit price* due to the the availability of the smaller 40 fl oz size, enabling consumers to purchase at a lower cost per unit. Lastly, the Warehouse format provides the lowest *minimum volume price* by offering the largest package size (135 fl oz) of Seventh Generation, reducing the cost per volume relative to smaller packages available at other retailers.

3.2 Estimation

We estimate price indices separately for each category-year combination using the methodology detailed in full in Web Appendix C. For analysis, we center each of the price indices at

[8]For a consistent comparison, we treat products where the volume is measured in different units as different brands. For example, in the detergent category, we treat liquid-Purex and powdered-Purex as separate brands because volume prices are not comparable between the two.

Table 2: Sample Comparison of Laundry Detergent Availability and Pricing Across Different Stores (with Hypothetical Prices)

Product	**EDLP**	**HiLo**	**Dollar**	**Warehouse**
40 fl oz Purex	-	-	$7	-
84 fl oz Purex	$10	Regular $12, Sale $9	$11	$11
84 fl oz Seventh Generation	$15	-	-	-
135 fl oz Seventh Generation	-	-	-	$18
Relative Price	Lowest	Highest		
Min Monthly Price		Lowest	Highest	Highest
Min Unit Price		Highest	Lowest	
Min Volume Price		Highest		Lowest

the category-year level.

We construct two sets of price indices: (1) indices based solely on RMS data, where a store s is defined at the physical store level, as observed in the RMS data and (2) indices using both RMS and HMS data, where a store s is defined as retailer-county due to the lack of physical store identifiers in the HMS data. We primarily focus on the combined RMS-HMS indices as they provide more comprehensive coverage of households' choice sets. To assess robustness, we also replicate key analyses using RMS-only-based price indices and report these results when relevant.

3.3 Variation in Price Indices

This section examines the variation in our recovered price indices, highlighting three key observations that inform our analyses of retailer differentiation (Section 4) and consumer store visit decisions (Section 5).

First, our indices capture systematic differences in pricing strategies across major retailer formats: grocery, warehouse, discount, dollar, and drug. We also find meaningful variation in pricing strategies *within retailer format*. For example, we find that some grocery stores adopt pricing strategies that more closely mirror discount or dollar store retailers than some of their grocery retailer counterparts. Past literature examining access to fresh produce has typically focused on fresh assortment differences across retailer formats. Our findings suggest that access to fresh produce may be determined not only by its presence in a retailer's assortment but also through the retailer's pricing strategy. For the most constrained consumers, fresh produce may be effectively out of reach at HiLo retailers or those promoting volume discounts, even if technically available, due to their limited ability to buy in bulk or shift purchases over time. We investigate this relationship further in Section 5.

Second, we find that although pricing strategies are largely determined at the retailer level, they also vary substantially across retailer-product departments in ways that suggest strategic differentiation by retailers. To account for this, in our analysis of household retailer visits in Section 5, we construct household-specific evaluations of retailer pricing strategies, weighting each price index by the relative importance of the corresponding product department in the household's basket.[9]

Finally, consistent with research on uniform pricing (e.g., DellaVigna and Gentzkow (2019)), we find that retailers maintain remarkably consistent pricing strategies across geographic markets. This finding allows us to overcome data sparsity in the HMS data and recover more representative consumer choice sets (see Web Appendix C). Second, it justifies our focus on retailer-county rather than store-level variation in our analyses of differentiation (Section 4) and consumer retailer visits (Section 5).

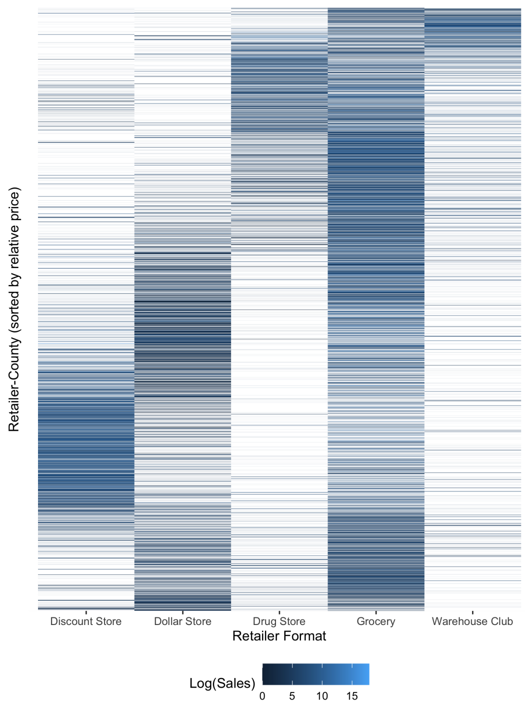
3.3.1 Retailer Price Indices

To analyze how pricing strategies vary across and within retail formats, we aggregate our four category-retailer-county-year level price indices $\hat{\mathbf{p}}_{csy}$ to the retailer-county-year level as sales-weighted averages. Figure 1 reveals distinct pricing strategy differences across formats. Warehouse clubs offer low volume prices, but accessing those savings involves lumpy spending, as products are sold in large quantities with higher unit prices. Dollar stores, in contrast, lead with the lowest unit prices, but are less attractive on the volume price, relative expensiveness and monthly discounts dimensions. Discount stores offer more favorable volume prices than dollar stores, but charge higher unit prices – requiring more lumpy spending – and offer minimal monthly discounts. Drug stores exemplify HiLo pricing by maintaining high relative and volume prices while featuring deep monthly discounts.[10] Although news coverage has highlighted specific pricing practices of certain formats – e.g., dollar stores' low unit prices, but poor

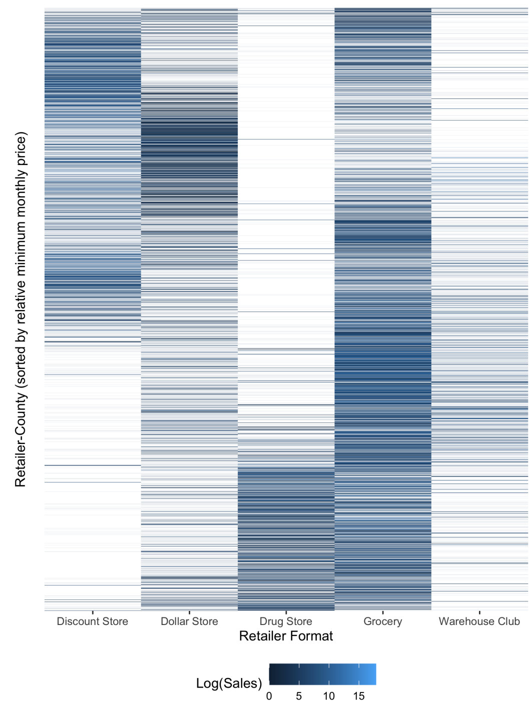
[9]This approach also aligns with Clerides et al. (2023)) who find that store expensiveness depends on a household's shopping basket.

[10]Figure 2 presents the non-residualized minimum monthly price, showing that drug store discounts are generally not deep enough to significantly lower their position in the price distribution as a group.

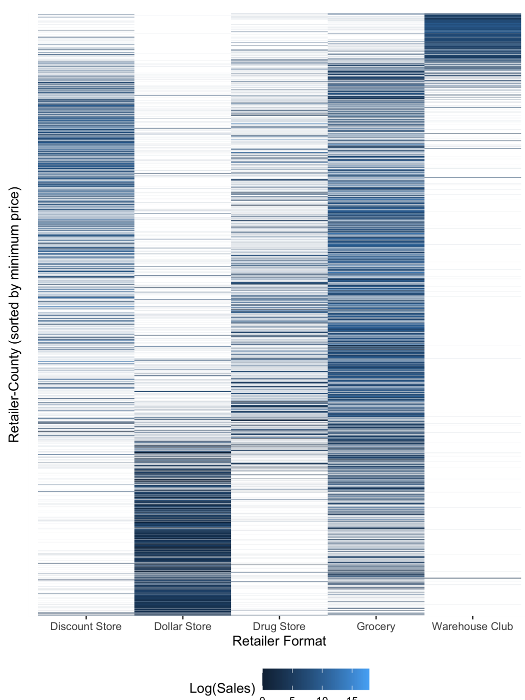
Figure 1: Price Indices by Retailer Format



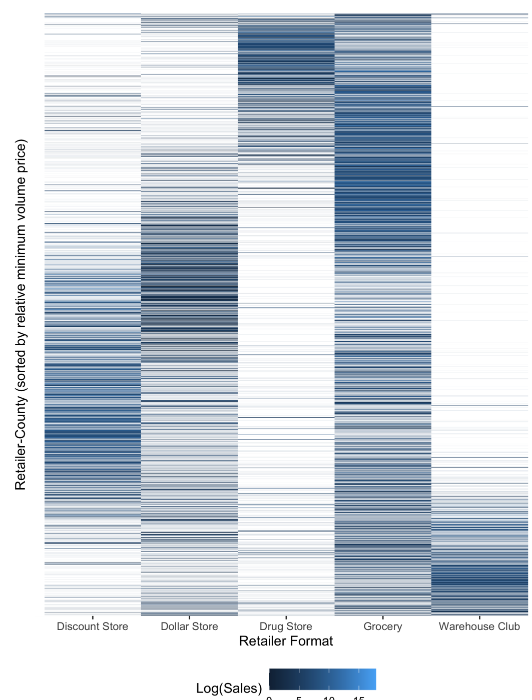
(a) Sorted by Relative Price



(b) Sorted by Minimum Monthly Price (Res)



(c) Sorted by Minimum Price



(d) Sorted by Minimum Volume Price

This figure illustrates how retailers position themselves in pricing strategy space across different retail formats. Each point represents a retailer-county-year observation in 2019, with pricing strategies calculated as sales-weighted averages of the centered retailer-county-category-year measures derived from estimation step 2 in Web Appendix C. The y-axis orders retailers (retailer-county) from lowest (bottom) to highest (top) price index values.

volume prices (e.g., Rogelberg (2024)) – our analysis provides the first systematic, data-driven documentation of pricing strategy differences across retail formats.

Notably, grocery retailers exhibit remarkable diversity in their pricing strategies, spanning much of the range observed across other retail formats. Some grocery retailers follow HiLo pricing, while others adopt EDLP or low unit price strategies more characteristic of discount or dollar stores, respectively.

3.3.2 Retailer-Department Price Indices

We next consider the differences in pricing strategies based on the types of products carried by retailers. To do so, we aggregate our four category-retailer-county-year level price indices $\hat{\mathbf{p}}_{csy}$ to the retailer-county-*product department*-year level as sales-weighted averages.[11]

We examine the variation in the product-department-specific price indices by estimating a sales-weighted regression of each of the indices on incremental fixed effects and reporting the adjusted R^2 from these regressions.[12] Table 3 confirms the patterns highlighted in Figure 1. Specifically, retailer format-state-product department fixed effects (first highlighted row in each set of regressions) explain a substantial share of variation in price indices, particularly for the unit price measure. The inclusion of retailer fixed effects (second set of highlighted rows) further increases explanatory power, indicating that pricing strategies are largely determined at the retailer level. Finally, although adding product department interactions provides a smaller incremental improvement (third set of highlighted rows), it accounts for much of the remaining variation, especially for the minimum monthly price measure. This suggests that product needs may play a meaningful role in shaping how households assess retailer pricing strategies, motivating our construction of household-specific price indices in Section 5.

Notably, product department alone has limited explanatory power (first two non-highlighted rows in each set), implying that pricing strategies are not inherently tied to product types,

[11]In what follows, we use retailer-county-department- rather than store-category-level price indices for two key reasons: (1) it allows us to incorporate HMS-only retailers, thereby better reflecting the consumer choice set and (2) department-level aggregation reduces noise in the household-specific evaluations of the price indices in Section 5. Table D.3 presents parallel analyses using the most granular price index measures based solely on RMS data (estimation step 1), with results that align with the more aggregated department-level findings.

[12]We weigh the regressions by total store-category-year revenues, deflated to January 2010 dollars using the U.S. Bureau of Labor Statistics (2024) consumer price index. This weighting accounts for the large dispersion in annual revenue volume across the different retailers and product departments. Nevertheless, the substantive results remain unchanged without weighting.

Table 3: Adjusted R^2: Regression of Price Indices on Fixed Effects (2012-2020)

Fixed Effects	Rel Price (1)	Min Price (2)	Min Month Price (Res) (3)	Vol Price (4)
RMS Only				
Dept × Year	0.03	0.09	0.04	0.04
Dept × Year × State	0.24	0.25	0.19	0.17
Dept × Year × State × Format	0.79	0.86	0.50	0.71
Dept × Year × State × Format + Retailer × Year	0.94	0.95	0.79	0.88
Dept × Year × State × Format + Retailer × Year × State	0.94	0.95	0.80	0.89
Dept × Year × State × Format + Dept × Retailer × Year	0.97	0.97	0.94	0.95
Dept × Year × State × Retailer	0.98	0.97	0.95	0.96
HMS and RMS				
Dept × Year	0.05	0.09	0.03	0.05
Dept × Year × State	0.30	0.21	0.23	0.23
Dept × Year × State × Format	0.68	0.74	0.52	0.64
Dept × Year × State × Format + Retailer × Year	0.84	0.83	0.81	0.82
Dept × Year × State × Format + Retailer × Year × State	0.85	0.84	0.82	0.83
Dept × Year × State × Format + Dept × Retailer × Year	0.89	0.87	0.91	0.87
Dept × Year × State × Retailer	0.90	0.88	0.93	0.89

This table presents the adjusted R^2 from regressions of the price indices on varying levels of fixed effects, weighted by total deflated revenue. We present results for both the RMS-only price indices (Step 1 in Web Appendix C) and the price indices constructed using both RMS and HMS data (Step 2 in Web Appendix C).

but rather reflect strategic choices by the retailers. Additionally, geographic variation within retailer contributes little to explaining differences in price indices (last two non-highlighted rows), consistent with prior research on uniform and zone pricing (Adams and Williams (2019), DellaVigna and Gentzkow (2019), Hitsch et al. (2021)).

Building on the findings that pricing strategies primarily vary at the retailer and retailer-product department level, we next provide suggestive evidence that a retailer's pricing strategy is part of a broader differentiation approach – one that entails costly co-investments in product assortment and store location decisions.

4 Evidence of Costly and Targeted Differentiation

In this section, we examine how differences in retailer pricing strategies are accompanied by complementary assortment and retailer location differences, providing evidence that pricing strategies are co-determined with other fixed-cost investments as part of a broader differentiation strategy.

Existing research on grocery retail pricing has highlighted the role of assortment and retailer quality in sustaining distinct pricing strategies and targeting specific consumer segments (e.g., Lal and Rao (1997), Ellickson (2006)). Other studies have also shown that certain retailer formats, such as HiLo retailers or dollar stores, strategically locate near their target segments (e.g., Ellickson and Misra (2008), Caoui et al. (2022)). This prior work has focused on specific pricing strategies, price levels, or retailer formats. We build on this work by documenting the costly differentiation – both in assortment and location – that sustains the full range of pricing strategies across all major retailer formats in the market. These patterns underscore that a retailer's choice of pricing strategy is inherently tied to its broader differentiation decisions, which, in turn, shapes the trade-offs faced by the most constrained households.

4.1 Assortment Differentiation

We first document the differences in assortment quality and variety that underlie variation in retailer pricing strategies and point to differences in fixed-cost investments across retailers. Table 3 shows that a large share of the variation in price indices is explained by retailer format-state-product department fixed effects. This pattern suggests that pricing strategy differences may result from the different costs of carrying different product types in different geographies (e.g., due to transportation costs) or constraints of specific retail formats (e.g., the limited footprint of dollar stores restricting their ability to carry products requiring specialized space). Given that most of the remaining variation in pricing strategies occurs across retailers and retailer–product departments, we next construct measures of assortment quality and variety at these levels to examine their relationship with pricing strategies.

At the retailer level, we assess assortment quality and variety through the presence of specific product departments and private label offerings. We focus on product departments that may incur higher operational costs – such as those requiring regulatory compliance (e.g., alcoholic beverages), specialized space (e.g., deli), refrigeration or freezing capabilities (e.g., milk, meat, frozen foods, fresh produce), or complex inventory management (e.g., fresh produce) – while also serving as a point of differentiation for consumers. Within fresh produce, we distinguish between barcoded (SKU) items and random-weight items, which may vary in quality and impose different inventory management costs. We also consider the presence of private label products, which require upfront investment but can provide retailers with greater pricing flexibility, potentially enabling lower prices (Kapner (2019)).

Table ?? shows that retailers carrying specialty product departments tend to have higher prices, suggesting that these departments may reflect higher operating costs, serve as differentiation points, or both.[13] The negative relationship between minimum monthly price and private label presence suggests that private labels may serve to enable deeper discounts at retailers that carry them. Notably, retailers offering random-weight produce tend to have higher relative prices and are more likely to use HiLo pricing – defined as a retailer-product department with above-zero relative prices and below-zero minimum monthly prices – potentially making them less appealing to consumers with tight, fixed grocery basket budgets. On the other hand, although retailers carrying barcoded produce offer more stable pricing, their higher minimum unit prices may still be out of reach for the most constrained households.

At the retailer-product department level, we measure assortment depth and quality by examining the number of categories within a department, the number of distinct brands and sizes in the average category in the product department, median name brand quality and package size differentiation. To capture median brand quality, we use the brand fixed effects $\hat{\xi}_b$ estimated using equation 4 as proxies for brand quality and compute the median value among name-brand products within each retailer-department. To capture package size differentiation, we calculate the share of UPCs within each retailer-product category that fall into the top or bottom quartile of available sizes across all retailers. We then average these shares across categories to obtain retailer-product department measures. These two metrics measure whether a retailer's offerings are skewed toward larger or smaller package sizes, which may be differently costly for retailers to provide and hold differential appeal to more or less constrained households.

Table ?? shows that retailers carrying higher-quality brands tend to have higher relative, volume, and especially unit prices, but not systematically greater price variability, whereas retailers offering greater brand variety are more likely to adopt HiLo pricing. Retailers with broader size assortments – and to some extent, category assortments – are more likely to follow everyday low pricing, characterized by lower relative prices, higher minimum monthly

[13] In the main text, we focus on all in-sample retailers. In Table E.4, we show that this result is not driven exclusively by differences across retailer formats by documenting similar relationships for grocery retailers only.

prices, and lower unit prices. These patterns suggest that retailers with reliably low prices may achieve cost efficiencies by offering a one-stop shopping experience with broad size and category variety, but limited, lower-end brand selection. In contrast, stores offering wider brand variety and higher-quality brands may sustain higher prices.

Additionally, retailers emphasizing larger package sizes tend to offer lower and more stable prices, both in relative and volume terms, while those focused on smaller package sizes provide lower unit and relative prices but shallower monthly discounts. The strongest relationship is between large package sizes and low volume prices, underscoring the substantial savings at warehouse-club-style retailers. These relationships highlight the joint pricing-package size differentiation with varying appeal to different consumer segments: (1) moderate package sizes and EDLP pricing may hold the most appeal for larger lower-income households, (2) bulk buys with low volume prices may attract larger higher-income households, and (3) small package sizes with low minimum prices may attract the most constrained low-income households that cannot take advantage of intertemporal or volume discounts.

Overall, these findings support the existence of costly assortment differentiation that reinforces pricing strategy differences and, in conjunction with those strategies, is tailored to specific consumer segments.

4.2 Geographic Differentiation

Sustaining particular pricing and assortment strategies may require sufficient scale and throughput, which retailers may achieve by tailoring their offerings to local demand conditions. To capture these co-investments in retailer location, we next examine retailers' joint location and pricing strategies, focusing on how pricing strategies align with local market characteristics.

To capture the localized nature of grocery retail, we construct ZIP-code specific retailer consideration sets as the complete list of retailers (both RMS participants and those observed only in HMS) visited by any household residing in the same ZIP code during that year.[14] We use these same consideration sets in Section 5 to analyze how household constraints shape retailer selection and shopping behavior.

For each retailer-ZIP pair, we track the retailers' pricing strategies, the demographics of

[14]Retailers generating less than 1% of their total revenue from a given ZIP code are excluded from that ZIP code's consideration set.

Table 4: Pricing Strategies and Local Household Income

Model:	Rel Price (1)	Min Month Price (Res) (2)	Min Price (3)	Vol Price (4)
Zip Median Income ($10,000)	0.08***	-0.009***	0.23***	0.02***
	(0.003)	(0.0009)	(0.009)	(0.005)
County Median Income ($10,000)	0.02***	-0.0008	0.05***	0.01**
	(0.004)	(0.0007)	(0.010)	(0.005)
Fixed-effects				
Year-State	Yes	Yes	Yes	Yes

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

This table shows that retailers make joint location-pricing strategy decisions that are targeted towards the local (ZIP) income. Both county median income and county household size is calculated as the weighted average across all ZIP Codes in the county, excluding the focal ZIP Code.

households residing in the ZIP code and the demographics of surrounding ZIPs in the same county. We obtain ZIP code-level demographic data from the American Community Survey (ACS) 5-Year Estimates for the U.S. Census Bureau's ZIP Code Tabulation Areas (ZCTAs) for 2015 and 2020 (United States Census Bureau (2020)) and focus our analysis on these two years in which the demographic data are available. To derive the demographics of other ZIP codes within the county, we compute a population-weighted average of median incomes across all other ZIP codes. The contrast between the local demographics and the demographics of the broader geographic area allows us to examine the extent to which these pricing strategies are locally targeted.

Table 4 shows a significant relationship between a ZIP code's median income and the price indices of available retailers. Specifically, retailers in higher-income ZIP codes tend to have higher relative prices, greater price variation (minimum monthly price), and higher unit and volume prices. These findings provide direct evidence that retailers with different pricing strategies locate near and cater to distinct consumer segments. Notably, the income of surrounding ZIP codes within the county exhibits similar but weaker relationships with price indices. This pattern suggests that while retailer co-location is influenced by broader market characteristics, it is also highly localized, with the immediate area's demographics playing a stronger role in shaping the retail landscape.

Table 5: Pricing Strategies and Local Household Demographics

Model:	Rel Price (1)	Min Month Price (Res) (2)	Min Price (3)	Vol Price (4)
Zip Median Income ($10,000)	0.04***	0.0009	0.14***	-0.01**
	(0.003)	(0.001)	(0.01)	(0.006)
% on Food Stamps	0.003	0.06***	-0.10	-0.13
	(0.06)	(0.02)	(0.18)	(0.09)
Employment Rate	0.51***	-0.14***	-0.68	1.1***
	(0.13)	(0.05)	(0.54)	(0.25)
Avg Household Size	-0.04***	-0.008***	-0.009	-0.06***
	(0.01)	(0.003)	(0.04)	(0.02)
% Single Head of Household	0.23***	-0.10***	0.14	0.32***
	(0.05)	(0.01)	(0.15)	(0.07)
% With Vehicle	-0.30***	-0.06***	0.86***	-0.77***
	(0.05)	(0.02)	(0.17)	(0.08)
log(Population Density)	0.02***	-0.010***	0.08***	-0.008**
	(0.002)	(0.0007)	(0.007)	(0.003)
log(Median Home Value)	0.16***	-0.04***	0.19***	0.19***
	(0.01)	(0.003)	(0.03)	(0.02)
% Single Family Homes	0.07***	-0.007*	-0.004	0.03
	(0.01)	(0.004)	(0.04)	(0.02)
Fixed-effects				
Year-State	Yes	Yes	Yes	Yes
Fit statistics				
Observations	27,902	27,902	27,902	27,902
R^2	0.40504	0.91946	0.26855	0.23794
Within R^2	0.12186	0.88100	0.09133	0.03499

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

This table shows that retailers make joint location-pricing strategy decisions that are targeted towards the local (ZIP) demographics.

Going beyond household income, Table 5 describes the relationship between retailer pricing strategies and additional demographic factors that correspond to varying consumer constraints. As before, we find that retailers with higher relative and unit prices are more prevalent in higher-income neighborhoods. Greater food stamp prevalence is linked to reduced price variation, suggesting that retailers in these areas offer more stable prices that may be particularly attractive to the most constrained households. Employment rate correlates positively with relative prices and price variation, with a particularly strong link to volume prices. This suggests that retailers requiring lumpy expenditures for bulk packaging are less likely to operate in high-unemployment areas, where households have limited ability to purchase in bulk.

Areas with larger families tend to have retailers with lower relative and volume prices, likely because these households can better leverage larger package sizes. In contrast, areas with more single-headed households – who typically have lower consumption and less time for comparison shopping – tend to have higher prevalence of higher relative prices, higher volume prices, and more price variation.

Vehicle ownership emerges as one of the most significant factors in our analysis. Areas with higher vehicle ownership show lower relative and volume prices, but higher price variation and unit prices. This aligns with expectations: vehicle access facilitates comparison shopping across multiple retailers, potentially driving down relative prices. It also reduces transportation costs, making it easier for consumers to take advantage of price variations and bulk purchases.[15]

Overall, these findings support the idea that retailers make joint location and pricing decisions to target distinct consumer segments. Supplementary analysis in Web Appendix E.2 further confirms these patterns, showing similar consumer sorting behavior and consistent targeting strategies among grocery retailers.

5 Household Constraints as Drivers of Consumer Choice

In the previous section, we provide evidence that distinct pricing strategies are accompanied by costly, targeted differentiation in assortment and location. Next, we analyze household retailer choice and shopping behavior to demonstrate that consumer decisions are substantially shaped by their budget constraints.

We first focus on household decisions of which retailers to visit, showing that, consistent with households being substantially influenced by their budget constraints, lower-income and larger households are more likely to visit retailers that offer low and *stable* prices. We then examine savings foregone within retailers, showing that while lower-income and larger households leave the least brand savings on the table, they forgo the most intertemporal savings and among the highest volume savings, suggesting that their grocery budgets bind at the shopping trip level.

[15] The effects of population density are less straightforward. While denser areas might support more retailers and thus foster competition, they may also be more challenging to navigate, limiting consumers' willingness to travel. Our analysis shows only modest effects of density on retailer pricing strategies. Similarly, higher median home values could indicate either greater disposable income or larger storage spaces. Our results suggest both interpretations may be valid, as higher home values correlate positively with both relative prices and price variation. A higher proportion of single-family homes shows similar correlations.

5.1 Retailer Visits

5.1.1 Specification and Identifying Variation

To examine how retailer pricing strategies influence consumer choice of retailer, we supplement the retailer price indices, assortment and geographic location data from Sections 3 and 4 with shopping trip records from the NielsenIQ HMS data. These data capture the retailer visited and the date of each trip for every household in the panel. Using consumer choice sets defined in Section 4.2 and focusing on consumer shopping trips in 2014-2020, our estimation sample includes 75,849 consumers making 32,038,019 grocery visits.

Our findings in Section 3.3.2 indicate meaningful variation in price indices across product departments within the same retailer, implying that households with different product needs may perceive a retailer's pricing strategy differently. To account for these differences, we construct consumer-specific evaluations of the four price indices. We operationalize this by first aggregating retailer-category measures to the retailer-department level and then weighting them based on each consumer's spending in those departments within a given year. Let $\hat{\mathbf{p}}_{cry}$ represent one of the four price indices measured at the category-retailer-state-year level.[16] Then we create individual-specific $\hat{\mathbf{p}}_{iry}$ as follows:

$$\hat{\mathbf{p}}_{dry} = \frac{\sum_{c \in C_d} \hat{\mathbf{p}}_{cry} \times \text{sales}_{cy}}{\sum_{c \in C_d} \text{sales}_{cy}} \tag{5}$$

$$\hat{\mathbf{p}}_{iry} = \frac{\sum_{d \in D} \hat{\mathbf{p}}_{dry} \times \text{sales}_{idy}}{\sum_{d \in D} \times \text{sales}_{idy}} \tag{6}$$

Following the same approach, we create household-specific evaluations of a retailer's assortment $\hat{\mathbf{a}}_{iry}$, by weighting the product department-retailer-state-year variables from Section 4 according to each household's spending in different departments within a given year.

We next assess how household characteristics and retailer pricing strategies influence household choice of which retailers to visit. Specifically, focusing on ZIP-level choice sets, we estimate a Poisson regression of the number of household visits to a retailer on the retailer's price in-

[16] Our findings in Sections 3 and 4 guide our decision to use retailer-state rather than retailer-county variation. We observe minimal geographic variation in retailer pricing strategies but retain state-level differentiation to account for potential differences in assortment. These differences may arise from factors such as varying costs or regulatory requirements for specific product categories (e.g., alcoholic beverages), which could contribute to broader retailer differentiation.

dices, assortments, and their interactions with household demographics. Letting v_{irt} denote the number of visits by household i at retailer store r in week t:

$$\log(E(v_{irt})) = \alpha_D D_{iy} + \sum_{p \in P} (\beta^p + \beta^p_D D_{iy}) \hat{\mathbf{p}}^p_{iry} + \sum_{a \in A} (\delta^a + \delta^a_D D_{iy}) \hat{\mathbf{a}}^a_{iry} + \gamma_{zrym} \tag{7}$$

where D_{iy} represents the demographic characteristics of household i in year y, while $\hat{\mathbf{p}}_{iry}$ and $\hat{\mathbf{a}}_{iry}$ represent household i's evaluation of retailer pricing strategies and assortments (as defined in equation 6). In the main specification, D_{iy} includes household income (described in Section 2) and household size (from the NielsenIQ HMS data), which together capture the primary sources of variation in household constraints.[17] γ_{zrym} accounts for fixed effects at the ZIP-retailer-year-month level, and $E(v_{irt})$ is assumed to follow a Poisson distribution. The resulting coefficients and interaction terms can be interpreted as changes to the elasticity of retailer visits to the pricing indices.

Our data and approach allow us to control for the other dimensions of retailer differentiation in Section 4 to zero in on the role of pricing strategies on consumer retailer visits. In particular, the assortment controls capture consumer evaluation of the differences in assortments at retailers with different pricing strategies,[18] and the ZIP-retailer choice sets condition on retailers' geographic location choices. The granular ZIP-retailer-year-month fixed effects γ_{zrym} further allow us to control for a wide range of potential confounding factors. These fixed effects capture local competition, the regional popularity of a retailer, the accessibility of the retailer through cars and public transit, seasonal preferences or promotions, and macroeconomic conditions. The parameters are primarily identified by comparing how consumers with different demographics, residing in the same ZIP code during the same month, choose to visit retailers with varying pricing strategies in the product departments most relevant to their needs.

5.1.2 Results

Because our specification includes multiple, correlated pricing measures, the coefficient estimates must be interpreted carefully. For example, the coefficient on *minimum monthly price* reflects the effect of a change in *minimum monthly price* while holding *relative price* and other

[17] In Appendix F, we present results incorporating a broader set of household demographics.

[18] Past literature has shown that product assortment, including prevalence of private label, can impact consumer store choice and loyalty (Sudhir and Talukdar (2004), Briesch et al. (2009), Seenivasan et al. (2016))

factors constant. Thus, an increase in *minimum monthly price* – with average prices fixed – implies a decrease in price variation, as prices become more tightly centered around the mean. For instance, if a product has an average price of $2 and is priced at $2 every week, the *minimum monthly price* equals the average. In contrast, if the price is $1 in two weeks and $3 in the other two, the *minimum monthly price* falls to 50% below the average, even though the average remains $2. In this case, a negative coefficient on *minimum monthly price* indicates that households disproportionately avoid retailers with greater price variation. Applying this logic, after accounting for *relative price*, a negative coefficient on *minimum unit price* suggests that households value the availability of products with very low unit prices, while a negative coefficient on *minimum volume price* indicates a preference for volume savings.

Notably, the estimates $\hat{\beta}^p$ and $\hat{\beta}^p_D$ in our full specification (equation 7) capture the appeal of pricing strategies themselves. Because the specification controls for both the local appeal of specific retailers and household evaluations of assortment features, these coefficients reflect households' visit elasticities to deviations in pricing strategies, conditional on broader retailer investments in assortment and location. Put differently, after accounting for assortment and location, we interpret these coefficients as revealing preferences for pricing strategies per se, net of preferences for the retailer or its products.

Table 6 presents our main findings. Column 1 demonstrates how retailer visit frequency responds to changes in a retailer's *relative price* and how this relationship varies with income and household size.[19] We illustrate these relationships with two representative households: a "constrained household" with an annual income of $20,000 and four members, and an "unconstrained household" with an annual income of $100,000 and two members. According to the baseline results in column (1) of Table 6, a 10% increase in relative prices will reduce the visit frequency by 4.7% for the constrained household, and increase visit frequency 6.5% for the unconstrained household.

In column (2), we introduce controls for household-specific evaluations of retailer assortment and their interactions with household demographics. Our adjusted model indicates that a 10% increase in *relative price* will reduce the visit frequency by 2.0% for the constrained household, and increase visit frequency 5.8% for the unconstrained household.

Column (3) incorporates our comprehensive set of four pricing metrics, showing that each

[19]In Web Appendix F, we present the results for the coefficients of all assortment variables and interactions.

Table 6: Retailer Visits, Pricing Strategies, and Household Characteristics

	log(E(Weeks Visited in Month))		
Model:	(1)	(2)	(3)
Variables			
Income ($10,000)	0.007***	-0.0001	-0.004
	(0.0001)	(0.006)	(0.006)
Household Size	0.010***	0.13***	0.14***
	(0.0004)	(0.02)	(0.02)
Relative Price	-0.39***	0.28***	1.4***
	(0.01)	(0.01)	(0.07)
Min Monthly Price			-1.1***
			(0.07)
Min Unit Price			0.03**
			(0.01)
Min Vol Price			-0.08***
			(0.02)
Income ($10,000) × Relative Price	0.12***	0.06***	0.15***
	(0.001)	(0.001)	(0.004)
Income ($10,000) × Min Monthly Price			-0.17***
			(0.004)
Income ($10,000) × Min Unit Price			0.06***
			(0.001)
Income ($10,000) × Min Vol Price			0.02***
			(0.002)
Household Size × Relative Price	-0.08***	-0.15***	-0.52***
	(0.003)	(0.004)	(0.01)
Household Size × Min Monthly Price			0.41***
			(0.01)
Household Size × Min Unit Price			0.03***
			(0.003)
Household Size × Min Vol Price			-0.03***
			(0.005)
Assortment Controls	No	Yes	Yes
Fixed-effects			
Year-Month-ZIP Code-Retailer	Yes	Yes	Yes
Fit statistics			
Observations	26,448,961	26,195,119	26,195,119
Pseudo R^2	-0.03216	-0.03013	-0.02956

Clustered (Year-Month-ZIP Code-Retailer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

This table presents the results of Poisson regressions analyzing the relationship between household visit frequency to a retailer per month and retailer pricing strategies, along with consumer demographics. The full specification in Column 3 follows equation 7.

has a distinct effect on retailer visits, and that these effects vary systematically across households of different demographics. As before, an increase in *relative price* similarly decreases visit frequency for the constrained household, but increases visit frequency for the unconstrained household (-3.8% and 18.6%, respectively). The coefficients on *minimum monthly price* reveal how households respond to a decrease in price variation. If the *minimum monthly price* decreases by 10% while holding *relative price* fixed, the increase in price variation results in the constrained household visiting 2.0% less frequently, while the unconstrained household visits 19.8% more frequently. This effect reflects two dynamics: First, lower-income households demonstrate a stronger aversion to price variation, with each $10,000 increase in income reducing the elasticity of retailer visits to *minimum monthly price* by 0.17. Second, each additional household member increases this elasticity by 0.41, likely because larger households can more effectively utilize stockpiled products purchased during promotions.

Minimum unit price exhibits a distinct pattern. Holding other pricing metrics and assortment fixed (e.g., broad measures of available package sizes), a decrease in minimum prices indicates presence of products of very small package sizes or low-quality brands in the assortment. We find that this reduces the visit frequency of both the constrained and unconstrained household; however, while the constrained household responds to a 10% decrease in *minimum unit price* by visiting 2.7% less frequently, the unconstrained household shows a stronger response, visiting 6.9% less frequently. We interpret this to mean that, although very low unit prices are universally less appealing, the ability to purchase at a very low unit price is disproportionately more useful to a household with a tight shopping trip budget constraint and less useful for households with unconstrained budgets.

Conversely, lower *minimum volume price* appeals more strongly to constrained households. Holding other pricing metrics and broad assortment features (e.g., brand quality) constant, a 10% reduction in *minimum volume price* increases visit frequency by 1.6% for the constrained household, yet reduces visit frequency by 0.6% for the unconstrained household. This contrast illustrates how households with more members can better capitalize on volume-based pricing strategies, likely due to their greater consumption potential.

Taken together, this analysis reveals distinct differences in the appeal of pricing strategies across households of different demographics. Lower income and larger households disproportionately favor low and *stable* pricing, while higher-income and smaller households are more drawn to on average higher but more variable-pricing. These patterns persist even after controlling for retailer assortments and local choice sets, suggesting that differences in preferences or proximity alone do not explain the results. In other words, because our full specification accounts for assortment and location, the observed responsiveness appears driven by consumers' reactions to pricing itself rather than to correlated retailer features generated by costly co-investments on the part of the retailers. Given that lower income and larger households are also typically more price-sensitive, these findings indicate that constraints – not lack of price sensitivity – limit these households' ability to benefit from certain discounts, making retailers that rely on such pricing strategies less attractive to them.

5.1.3 Additional Specifications and Robustness Checks

Appendix F presents several supplementary analyses to ensure robustness of these results. First (Appendix F.1), we incorporate a more comprehensive set of household demographics D_{iy} available in the NielsenIQ data. The main results remain robust, and the analysis presents additional insights; namely, that single-family home residents place greater emphasis on volume pricing, likely due to increased storage capacity, while households with one non-working head react more strongly to the minimum monthly prices, likely due to greater time flexibility to seek out better sales.

Second (Appendix F.3), to address potential measurement error – particularly for retailers observed only in the HMS data – we replicate our analysis using only RMS-participating retailers, whose price indices are constructed from the dense pricing panel available in the RMS data. The results remain qualitatively similar to our main findings.

Third (Appendix F.5), we estimate a discrete choice model, which requires higher-level data aggregation and allows for fewer controls and demographic interactions, but in return allows us to better capture household substitution across retailers with different pricing strategies. The results are qualitatively similar to our main specification.

Finally (Appendix F.4), we investigate whether differential transportation access across income groups might drive our results. While we cannot directly observe household-level transportation access, we incorporate ZIP code-level census data from 2015 and 2020, including per capita vehicle ownership rates, average work commute times, and public transit usage percentages. We then conduct separate analyses for ZIP codes above and below median values for these transportation metrics. These results are largely qualitatively similar across the different sets of ZIP codes.

This latter analysis suggests that transportation is not the primary source of constraints guiding consumer retailer visits decisions. In the next section, we examine the types of savings forgone within retailer to highlight that budget constraints that bind at the shopping trip level are the primary source of household constraints.

5.2 Foregone Savings

Next, we examine household shopping behavior within retailers and how its variation across household demographics reflects differences in constraints. Specifically, conditioning on retailer visited and grocery basket purchase, we examine households' ability to capitalize on three types of monetary savings: (1) intertemporal, (2) brand switching, (3) volume discounts, and (4) retailer switching. Our analysis provides empirical evidence of static budget constraints that bind at the shopping trip level among the lowest-income and largest households, rationalizing their preference for retailers with low and stable prices.

5.2.1 Counterfactual Prices and Foregone Basket Savings

We analyze shopping behavior using NielsenIQ HMS shopping baskets ("purchases" data). These data capture all UPCs purchased by a household at a particular retail chain on a given date. To ensure the analysis is restricted to shopping trips with consistently available information on both realized and "counterfactual" prices, we focus on shopping baskets where at least one item falls within the in-sample product categories at RMS-participant retailers only, for which we observe a dense price panel unavailable for HMS-only retailers. Our analysis conditions on realized shopping baskets and prices paid, as recorded in the RMS data. Web Appendix G provides additional details on the selection criteria and data preparation for this analysis.

We next calculate the potential savings a household could have achieved if they had purchased each of the products in their shopping basket at a counterfactual lower price. Let p^r_{jst} represent the realized (paid) price, p^f_{jst} the counterfactual (forgone) price, and q_{jst} the quantity of product j purchased. Conditioning on the products in the shopping basket, we express the monetary savings foregone in household i's shopping basket at a given retailer s and time t as the following percentage:

$$\text{foregone basket savings}_{ist} = 100\% * \left(1 - \left(\frac{\sum_{j \in \text{basket}_i} q_{jst} p^f_{jst}}{\sum_{j \in \text{basket}_i} q_{jst} p^r_{jst}}\right)\right) \tag{8}$$

We consider four different types of potential savings that a household may forgo in their shopping basket. First are inter-temporal savings that arise from purchasing a product in a week when it is not at its deepest monthly discount at the given retailer. The counterfactual price we use to compute this type of foregone savings is the lowest shelf price of a given product (UPC) at the same retailer within the purchase month. The foregone inter-temporal savings tells us what percentage of the basket purchase spend a household could have saved if they purchased *each product* in the basket at its lowest price available in that month.

Second, we consider brand savings that arise from purchasing a lower-priced brand. We compute the counterfactual brand price as the price of the lowest-priced alternative with the same package size but a different brand descriptor within the category-week. The foregone brand savings reflect the percentage of the basket purchase volume a household could have saved by purchasing the cheapest available brand offering the same package size.

Next, we consider volume savings that arise from volume discounts. We compute the counterfactual volume prices as the volume price a household would have paid had they purchased the package size most closely resembling the volume of their monthly consumption of the brand within the year. In contrast to the previous two savings measures, to compute the foregone volume savings, we use realized and counterfactual *volume* prices and total *volume* of the monthly consumption amount.

Lastly, we examine potential savings from switching from the focal retailer to the set of stores that offer the lowest price for some product in the shopping basket. We compute the counterfactual store prices as the price a household could have paid for the same product (UPC) at the retailer offering the lowest price in the given week. For each shopping basket, we limit the comparison set of retailers to those within the same county as the store where the basket was originally purchased.[20] The foregone store savings represents the percentage of the basket's purchase spend a household could have saved by visiting the fewest number of stores necessary

[20]One limitation of this foregone savings calculation is that our approach to computing counterfactual prices requires a dense panel of counterfactual prices that the household could have paid, limiting the analysis to the NielsenIQ RMS-participating stores that a household could have shopped at. In reality, households may have other non-RMS-participant retailers in the choice set, which we cannot account for in this analysis.

to secure the lowest price for each purchased product. Unlike past research that considers store savings from switching to the lowest priced store for the given shopping basket (Clerides et al. (2023)), this approach reflects potential savings from visiting multiple retailers to achieve the lowest basket price. Any variation in these savings across household incomes is likely to reflect differences in travel costs or the opportunity costs of time associated with visiting more stores in a given week.

5.2.2 Forgone Savings as Evidence of Consumer Constraints

Next, we examine how forgone savings vary across households by regressing forgone savings on two key household characteristics – income and size – which reflects differences in financial and non-financial constraints. To account for the household selection into retailers illustrated in Section 5, we include ZIP-retailer-year-month fixed effects. This allows us to interpret differences in forgone savings as reflecting how households of different income levels and sizes, within the same retailer-ZIP code, leave varying amounts of savings on the table, conditional on their actual shopping baskets.

Table 7 provides evidence that the lowest-income households (earning below $\$25,000$ annually) and the largest households (four or more members) face the greatest constraints in utilizing inter-temporal and volume discounts. Despite lower-income households not differing in the number of visits (column (3) of Table 6), columns (1) and (2) of Table 7 show that forgone inter-temporal savings decline with income, while column (2) further indicates that these forgone savings increase with household size. Column (6) suggests that these households are also substantially constrained in taking advantage of volume discounts, with only the highest two income quintiles leaving more potential savings untapped than the lowest-income households.

By contrast, foregone brand savings increase with household income and decrease with household size (columns (3) and (4)), suggesting that lower-income and larger households are more likely to trade-off brand name to achieve lower shopping basket costs. Finally, foregone retailer savings exhibit the smallest and least systematic differences across household incomes. Columns (7) and (8) suggest that while the lowest-income and largest households may achieve some additional savings by shopping at additional retailers with lower prices for their basket,

Table 7: Differences in Foregone Savings Potential by Household Income and Size

	Within Retailer Foregone Savings (%)						Across Retailer Foregone Savings (%)	
	Inter-temporal		Brand		Volume		Retailer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables: **Household Income**								
\$25,000-50,000	-0.10***	-0.17***	0.50***	0.68***	-0.13	-0.58***	-0.02	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)	(0.13)	(0.13)	(0.02)	(0.02)
\$50,000-80,000	-0.22***	-0.33***	0.84***	1.12***	0.25**	-0.48***	-0.07***	-0.09***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.12)	(0.13)	(0.02)	(0.02)
\$80,000-130,000	-0.57***	-0.71***	1.35***	1.74***	1.25***	0.25*	0.01	-0.02
	(0.03)	(0.03)	(0.04)	(0.04)	(0.13)	(0.13)	(0.02)	(0.02)
\$130,000+	-0.64***	-0.80***	2.21***	2.62***	2.08***	1.02***	0.16***	0.13***
	(0.03)	(0.04)	(0.05)	(0.05)	(0.15)	(0.16)	(0.03)	(0.03)
Variables: **Household Size**								
HH Size = 3		-0.17***		0.97***		-2.83***		0.00
		(0.02)		(0.03)		(0.11)		(0.02)
HH Size = 2		-0.30***		1.56***		-3.27***		-0.05***
		(0.02)		(0.03)		(0.09)		(0.01)
HH Size = 1		-0.54***		1.77***		-4.40***		-0.11***
		(0.03)		(0.04)		(0.11)		(0.02)
Fixed-effects								
ZIP-Retailer-Month	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	6,181,970	6,181,970	6,181,970	6,181,970	6,181,970	6,181,970	6,090,469	6,090,469

Clustered (ZIP-Retailer-Year-Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

This table presents the results of regressions of foregone savings on household chacacteristics indicative of consumer constraints: household income and size. Foregone savings are calculated as described in Section 5.2, and income quintiles are formed as described in Section 2.

Table 8: Foregone Savings Potential by Household Income

			Within Retailer Foregone Savings (%)			Across Retailer Foregone Savings (%)
HH Income	Avg. Monthly Baskets	Avg. Monthly Spend	Inter-temporal	Brand	Volume	Retailer
			(1)	(2)	(3)	(4)
≤25,000	8.6	243.8	10.8	20.1	9.3	6.9
\$25,000-50,000	8.6	271.6	10.8	20.6	9.4	6.8
\$50,000-80,000	8.3	296.7	10.5	21.1	9.4	7.0
\$80,000-130,000	8.2	318.3	10.1	21.9	9.5	7.6
\$130,000+	7.9	323.8	9.9	21.6	9.1	7.9

This table summarizes the percentage inter-temporal, brand, volume and retailer savings foregone by households of different incomes. Foregone savings are calculated as described in Section 5.2, and income quintiles are formed as described in Section 2.

the highest-income households stand to benefit the most from retailer switching.

The summary statistics of average foregone savings by household income in Table 8 provide insight into the magnitude of these potential savings. The table highlights that brand and inter-temporal savings are the most significant: households could reduce spending by 20–21% by switching to a different brand of the same package size (column (2)) and by 10–11% by purchasing products at their lowest monthly price (column (1)). Volume savings offer a comparable but slightly smaller potential, around 9% (column (3)). Finally, households could achieve 7–8% in savings on their realized basket by shopping at additional retailers within a given week.[21]

Importantly, we find that the lowest-income households forgo savings equal to a notable percentage of their income and at-home food budget. Focusing on savings induced by budget constraints, households in the lowest income quintile forgo \$316 annually due to limited ability to take advantage of inter-temporal discounts[22] and \$272 due to limited capacity to buy in bulk, each totaling approximately 1% of a \$25,000 annual income and 6-7% of the annual at-home food budget for this income group (Martin (2024)). Additionally, since these calculations are conditional on the retailers visited, the actual foregone inter-temporal savings may be understated, as lower-income households disproportionately shop at retailers with more stable and lower unit prices (Section 5).

Although the brand savings represent an even more substantial share of household income – \$588 annually or 2.3% of a $\$25,000$ income – realizing these savings would require switching to lower-priced brands, including private label products, potentially at the expense of quality.[23] This suggests that even the lowest-income households may be unwilling to trade off quality for lower prices. In contrast, inter-temporal and volume savings estimates are conditioned on UPC and brand, respectively, and therefore do not involve such quality trade-offs.

Taken together, these findings suggest that budget constraints significantly shape the savings different households can achieve and, consequently, the types of retailers they find most appealing. Specifically, higher-income households are better positioned than lower-income households to take advantage of inter-temporal discounts at HiLo retailers and quantity discounts at

[21] Table G.14 further details the changes required to realize these savings, including additional weeks and retailers visited, as well as the share of baskets involving a switch to private labels.

[22] Calculated by multiplying the foregone savings potential by the average annual spend across all baskets in Table 8 $= 0.108 \times 243.8 \times 12$.

[23] Table G.14 shows that achieving brand savings would require a switch to private label in approximately 17% and 19% of the lowest- and highest-income households' baskets, respectively.

those offering larger package sizes. Consistent with prior research showing that grocery store proximity has little impact on the healthfulness of consumer purchases (Allcott et al. (2019)), our results suggest that travel and time costs pose a less substantial barrier to achieving basket savings for the lowest-income households. However, we acknowledge that the measurement of forgone retailer savings may be influenced by incomplete capture of consumer choice sets in this analysis.

6 Implications for the Most Constrained Households

In Sections 3 and 4, we document how retailers differentiate their pricing strategies to target consumers with varying constraints, enabled by costly investments in assortment and location. In Section 5, we further show that retailer visit patterns and within-store forgone savings are consistent with budget constraints that bind at the shopping trip level for the lowest-income and largest households. These patterns reinforce the idea that retailer pricing strategies are designed to accommodate consumers facing differently binding constraints.

This section examines the implications of this retailer differentiation for the most constrained households. We first analyze how targeted pricing strategies affect the effective prices and product assortments available to these households. We then explore how retailer-driven geographic differentiation creates preference externalities that further impact their choice sets. Although a full welfare analysis is beyond the scope of this paper, our findings suggest that the fragmented grocery market, which caters to consumer constraints as well as preferences, imposes monetary costs and reduces access to higher quality healthful assortments for the most constrained households. Additionally, we find that constrained households in less disadvantaged areas face a trade-off between access to healthier retailer assortments and pricing strategies that are less aligned with their budget constraints.

6.1 Available Prices and Assortment

As shown in Section 5, the most constrained households disproportionately shop at retailers with low and stable prices due to their binding budget constraints. To assess the impact of this targeted differentiation, we compare assortments and *relative*, *minimum monthly*, and *volume*

Table 9: Price Indices of Retailers in the Bottom 20^{th} Percentile of Unit Prices

Price Index	25^{th} Perc.	Median	Mean	75^{th} Perc.
Min Price	9	14	13	17
Rel Price	12	29	35	50
Min Monthly Price (Res)	33	67	59	85
Vol Price	20	42	45	64

This table shows the distribution of price indices for retailers in the bottom 20^{th} percentile of minimum unit prices. Retailer is defined as retailer-county-year and retailers are assigned quintiles within county and year. The summary statistics for price indices are across all retailer-county-years.

Table 10: Fresh Produce Categories by Quintile of Unit Prices

Min Unit Price Quintile	25^{th} Perc.	Median	Mean	75^{th} Perc.
Q1 (20^{th} Perc. Unit Price)	0	0	7	20
Q2	0	10	11	22
Q3	0	21	14	23
Q4	10	21	16	23
Q5 (80^{th} Perc. Unit Price)	16	20	16	22

This table shows the distribution of the number of fresh produce categories by quintile of retailer unit price. Retailer is defined as retailer-county-year and assortments are computed at the retailer-state-year, as in Section 4.1. Fresh produce categories are the number of categories in the fresh produce (SKU) department at the retailer-state-year level.

prices at retailers in the bottom of the *minimum unit price* distribution – those disproportionally serving the most constrained consumers – to other retailers available in the same county and year.[24]

Table 9 shows that retailers in the bottom 20^{th} percentile of unit prices in their county rank much higher on all other pricing measures (*median* 29^{th}, 67^{th} and 42^{nd} percentiles for *relative*, *minimum monthly*, and *volume price*, respectively), suggesting that household constraints may cost the most constrained households lower prices in relative terms. In terms of healthful assortments, we find that only 48% and 63% of the lowest-unit-price retailers carry SKU and random-weight fresh produce, compared to 74% and 81%, respectively, for retailers in higher price tiers. Table 10 further highlights disparities in fresh produce variety, showing that the median retailer in the bottom quintile of unit prices carries zero fresh produce categories,

[24]This analysis focuses on *choice sets*, as fully understanding counterfactual household *choices* would require a micro-founded model of store selection and shopping behavior that accounts for consumer preferences and willingness to pay for healthful assortments.

whereas the median retailer in the top quintile carries 20.

These findings indicate that retailers offering low minimum prices – catering to budget-constrained consumers – do so at the expense of a deep and diverse fresh produce offering. As a result, even if households value or prefer healthful foods, their budget constraints may lead them to shop at stores with fewer and lower-quality options, ultimately reducing their likelihood of purchasing such products.

6.2 Externalities

Section 4.2 shows that retailers in lower-income ZIP codes tend to offer lower and more stable prices, consistent with the findings from Section 5, which indicate that lower-income and more budget constrained consumers are disproportionately drawn to such pricing strategies. These results highlight the presence of externalities in local retail markets, in which the predominant consumer group shapes the availability of retail options (Waldfogel (2003), Handbury (2021)). Consequently, individuals whose preferences and *constraints* differ from those of the surrounding community may have limited access to their preferred pricing strategies, potentially leaving certain consumer segments under-served.

This implies that the most constrained households in ZIP codes where the majority of residents share similar constraints are more likely to encounter retailers whose pricing strategies align with their needs. However, they also face the assortment trade-off described in Section 6.1. In contrast, the most constrained households living in areas with less budget-constrained neighbors encounter a different challenge. While they may have access to retailers offering a broader selection of healthier products, these assortments may be effectively out of reach due to the retailers' pricing strategies. As a result, their ultimate choice of retailer depends on the trade-off between adhering to a constrained budget and the difficulty of traveling to a retailer with more favorable pricing.

6.3 Discussion

Taken together, our results offer insights into the discussion on food deserts – areas where residents lack nearby access to grocery stores – and inform policies aimed at reducing the nutritional gap between the highest- and lowest-income households.

First, our research rationalizes food deserts as an equilibrium outcome shaped by constrained household choices and retailer pricing strategies that respond to these constraints. The historically dominant grocery retail format, characterized by frequent in-store price promotions (Ellickson et al. (2012)), enabled intertemporal and quantity discount price discrimination, where less price-sensitive consumers effectively cross-subsidized those more responsive to promotions and discounts. The entry of large discount stores and warehouse clubs into grocery retailing reshaped this landscape by targeting consumers willing to engage in lumpy spending in exchange for quantity discounts, fragmenting pricing and assortment strategies in the market (Ellickson et al. (2012) on Walmart entry; Courtemanche and Carden (2014) on warehouse club entry). At the same time, specialty retailers emerged, catering to consumers who prioritize specific product departments or assortments.[25]

This shift toward more refined targeting of less constrained households has reshaped the retail options accessible to those with the most constrained budgets. Although the traditional grocery format may not have fully met the needs of consumers least able to take advantage of intertemporal discounts, it likely offered lower prices than many of today's HiLo-pricing grocery retailers because it served a broader mix of shoppers, including price-sensitive consumers, bulk buyers, and quality-conscious customers.[26] However, as discounters and warehouse clubs drew these segments away, market fragmentation led to pricing strategies increasingly tailored to differently constrained consumers. This led to formats such as dollar stores (Caoui et al. (2022)) and retailers with stable and low minimum prices more generally (this study) increasingly targeting the most constrained.

Because pricing strategies and assortments are inherently linked, this fragmentation has left the most constrained households primarily reliant on retailers where affordability comes at the cost of nutritional variety and quality. Retailers offering the lowest prices to attract budget-constrained consumers often face trade-offs that restrict their ability to maintain a high-quality grocery assortment while keeping prices low. As a result, the most budget-friendly options come with a nutritional trade-off, while retailers that offer healthier, higher-quality assortments or volume discounts tend to be more expensive, targeting less constrained consumers who prioritize

[25]For a review of this evolution, see Ellickson (2016).

[26]Courtemanche and Carden (2014) show that retailers facing the strongest competition from warehouse clubs increased their prices.

quality or bulk savings, respectively.

Past research has examined food deserts as an equilibrium outcome, primarily attributing them to differences in consumer preferences for healthful food on the demand side (Allcott et al. (2019)). Our study adds nuance to this perspective by showing that consumers select retailers based on the overall affordability of their entire *shopping basket*, often trading off more healthful assortments to better meet their category needs within a constrained budget. In other words, our findings suggest that even if households share the same preferences for healthful food, differences in financial constraints may lead them to shop at retailers with varying levels of healthful assortments, ultimately influencing nutritional outcomes. However, quantifying the relative impact of constraints versus preferences is beyond the scope of this paper.

Our findings also have implications for policies aimed at reducing disparities in nutritional outcomes. While differences in nutrition education have been proposed as a contributing factor (Allcott et al. (2019)), our results suggest that education alone may be insufficient for the most constrained households. Policies that alleviate rigid grocery budget constraints may be more effective in enabling these households to access retailers with broader and more nutritious assortments. Healthful food subsidies, such as those examined in Olsho et al. (2016), Allcott et al. (2019), and Hinnosaar (2022), represent one such approach. Furthermore, our results suggest that subsidies may need to target the entire shopping basket at retailers with healthful assortments, as subsidizing a single product category may not be sufficient to draw in the most budget-constrained consumers.

Moreover, our results imply that policies that subsidize grocery store entry into former food deserts may need to consider not only whether a retailer improves access to produce but also the retailer's pricing strategy. Our findings suggest that if grocery retailers with relatively high or fluctuating prices enter areas with financially constrained households, they may fail to meaningfully improve access to fresh produce, even if they offer a wide selection.

7 Conclusion

In this paper, we examine pricing strategies in the retail market, documenting how fragmentation in joint pricing and assortment strategies has led to more precise targeting of distinct segments of less constrained consumers. As a result, the most constrained households are primarily served by low and stable price retailers that offer more limited depth fresh assortments. We highlight the implications of this market structure for the most constrained consumers and discuss policy considerations for reducing the nutritional gap between the lowest- and highest-income households.

References

Adams, B. and K. R. Williams (2019). Zone pricing in retail oligopoly. *American Economic Journal: Microeconomics 11*(1), 124–156.

Allcott, H., R. Diamond, J.-P. Dubé, J. Handbury, I. Rahkovsky, and M. Schnell (2019). Food deserts and the causes of nutritional inequality. *The Quarterly Journal of Economics 134*(4), 1793–1844.

Bell, D. R. and J. M. Lattin (1998). Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. *Marketing Science 17*(1), 66–88.

Briesch, R. A., P. K. Chintagunta, and E. J. Fox (2009). How does assortment affect grocery store choice? *Journal of Marketing research 46*(2), 176–189.

Bronnenberg, B. J., C. F. Mela, and W. Boulding (2006). The periodicity of pricing. *Journal of Marketing Research 43*(3), 477–493.

Cao, Y., J. A. Chevalier, J. Handbury, H. Parsley, and K. R. Williams (2024). Distributional impacts of the changing retail landscape. Working paper.

Caoui, E. H., B. Hollenbeck, and M. Osborne (2022). The impact of dollar store expansion on local market structure and food access. Available at SSRN 4163102.

Carlson, S., J. Llobrera, and B. Keith-Jennings (2021). More adequate SNAP benefits would help millions of participants better afford food. Accessed on Jul 1, 2024.

Child, P., T. Kilroy, and J. Naylor (2015). Modern grocery and the emerging-market consumer: A complicated courtship. Technical report, McKinsey & Company.

Chintagunta, P. K., J.-P. Dubé, and V. Singh (2003). Balancing profitability and customer welfare in a supermarket chain. *Quantitative Marketing and Economics 1*, 111–147.

Clerides, S., P. Courty, and Y. Ma (2023). Store expensiveness and consumer saving: Insights from a new decomposition of price dispersion. *Quantitative Marketing and Economics 21*(1), 65–94.

Courtemanche, C. and A. Carden (2014). Competing with Costco and Sam's Club: Warehouse club entry and grocery prices. *Southern Economic Journal 80*(3), 565–585.

DellaVigna, S. and M. Gentzkow (2019). Uniform pricing in US retail chains. *The Quarterly Journal of Economics 134*(4), 2011–2084.

Ellickson, P. B. (2006). Quality competition in retailing: A structural analysis. *International Journal of Industrial Organization 24*(3), 521–540.

Ellickson, P. B. (2016). The evolution of the supermarket industry: from A & P to Walmart. In *Handbook on the Economics of Retailing and Distribution*, pp. 368–391. Edward Elgar Publishing.

Ellickson, P. B. and S. Misra (2008). Supermarket pricing strategies. *Marketing Science 27*(5), 811–828.

Ellickson, P. B., S. Misra, and H. S. Nair (2012). Repositioning dynamics and pricing strategy. *Journal of Marketing Research 49*(6), 750–772.

Handbury, J. (2021). Are poor cities cheap for everyone? Non-homotheticity and the cost of living across US cities. *Econometrica 89*(6), 2679–2715.

Hinnosaar, M. (2022). The persistence of healthy behaviors in food purchasing. *Marketing Science 42*(3), 521–537.

Hitsch, G. J., A. Hortacsu, and X. Lin (2021). Prices and promotions in us retail markets. *Quantitative Marketing and Economics 19*, 1–80.

Kaplan, G. and G. Menzio (2015). The morphology of price dispersion. *International Economic Review 56*(4), 1165–1206.

Kaplan, G., G. Menzio, L. Rudanko, and N. Trachter (2019). Relative price dispersion: Evidence and theory. *American Economic Journal: Microeconomics 11*(3), 68–124.

Kapner, S. (2019, February). Kraft Heinz's tough approach on costs hurt grocery relationships. Date accessed: 2024-10-08.

Kelly, D. (2023, August). Why you'll almost never see name brands at Aldi. Date accessed: 2024-07-17.

Lal, R. and R. Rao (1997). Supermarket competition: The case of every day low pricing. *Marketing Science 1*, 60–80.

Martin, A. (2024). More adequate snap benefits would help millions of participants better afford food. Accessed on March 28, 2024.

Moshary, S., A. Tuchman, and N. Vajravelu (2023). Gender-based pricing in consumer packaged goods: A pink tax? *Marketing Science Articles in Advance*, 1–14.

Olsho, L. E., J. A. Klerman, P. E. Wilde, and S. Bartlett (2016). Financial incentives increase fruit and vegetable intake among supplemental nutrition assistance program participants: a randomized controlled trial of the usda healthy incentives pilot. *The American journal of clinical nutrition 104*(2), 423–435.

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

Palermo, A. (2024). Purex vs. Tide: Which laundry detergent is better? `https://prudentreviews.com/purex-vs-tide/`. Accessed: 2024-01-05.

Riss, C. (2024). Seventh Generation Vs. Tide laundry detergent. `https://livingproofmag.com/seventh-generation-vs-tide/`. Accessed: 2024-09-24.

Rogelberg, S. (2024, February). Meet the worst offenders on 'shrinkflation': dollar stores. Date accessed: 2024-07-17.

Seenivasan, S., K. Sudhir, and D. Talukdar (2016). Do store brands aid store loyalty? *Management Science 62*(3), 802–816.

Sudhir, K. and D. Talukdar (2004). Does store brand patronage improve store patronage? *Review of industrial organization 24*, 143–160.

Tax Policy Center (2024). Household income quintiles. `https://www.taxpolicycenter.org/statistics/household-income-quintiles`. Accessed: 2024-03-28.

Thomassen, Ø., H. Smith, S. Seiler, and P. Schiraldi (2017). Multi-category competition and market power: A model of supermarket pricing. *American Economic Review 107*(8), 2308–2351.

Traynor, S. (2024, April). Top 30 most expensive grocery stores in the world. Date accessed: 2024-07-17.

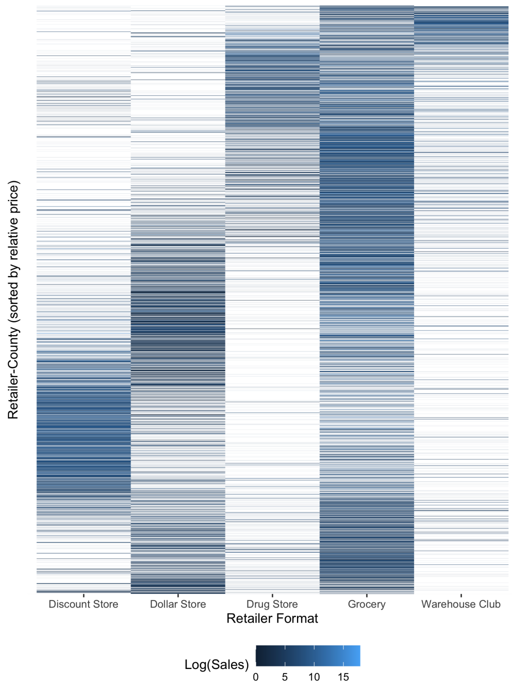
United States Census Bureau (2020). American community survey 5-year data (2009-2020). `https://www.census.gov/data/developers/data-sets/acs-5year.html`. Accessed: 2024-09-08.

U.S. Bureau of Labor Statistics (2024). Consumer Price Index for All Urban Consumers: All Items Less Food and Energy in U.S. City Average [CPILFESL]. `https://fred.stlouisfed.org/series/CPILFESL`. Accessed: 2024-01-17.

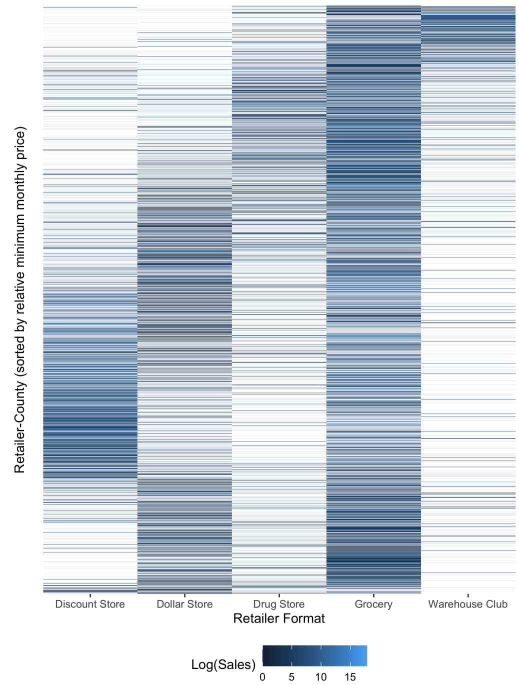
Waldfogel, J. (2003). Preference externalities: An empirical study of who benefits whom in differentiated-product markets. *RAND Journal of Economics 34*(3), 557–68.

Appendix

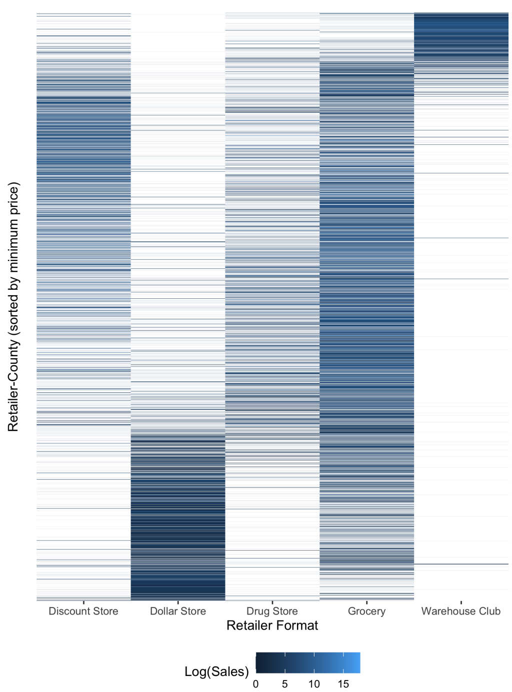
Figure 2: Price Indices by Retailer Format



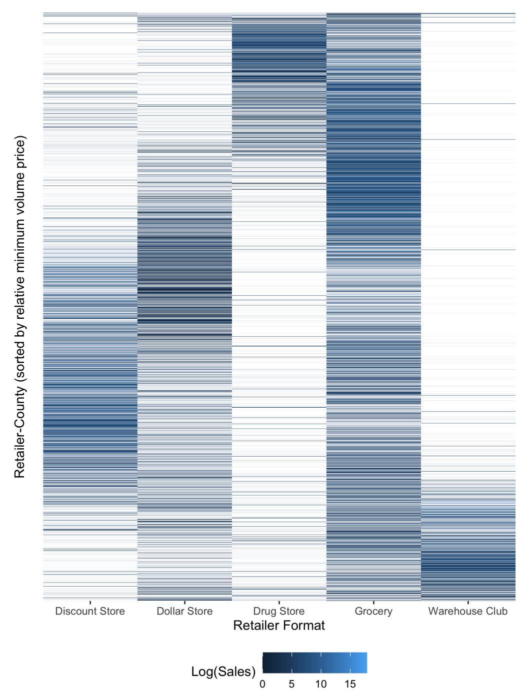
(a) Sorted by Relative Price



(b) Sorted by Minimum Monthly Price



(c) Sorted by Minimum Price



(d) Sorted by Minimum Volume Price

This figure illustrates how retailers position themselves in pricing strategy space across different retail formats. Each point represents a retailer-county-year observation in 2019, with pricing strategies calculated as sales-weighted averages of the centered retailer-county-category-year measures derived from estimation step 2 in Web Appendix C. The y-axis orders retailers (retailer-county) from lowest (bottom) to highest (top) price index values.

Dimensions of Retail Price Competition and Consumer Choice Constraints

Kristina Brecko Paul Ellickson Avery Haviv

May 7, 2025

Web Appendix

*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. The authors thank their colleagues at Rochester, seminar participants at UC Berkeley and UCLA Marketing Camp, as well as conference participants at the Choice Symposium and Marketing Science conferences for the helpful comments.

†Simon Business School at University of Rochester, kristina.brecko@simon.rochester.edu

‡Simon Business School at University of Rochester, paul.ellickson@simon.rochester.edu

§Simon Business School at University of Rochester, avery.haviv@simon.rochester.edu

A Household Trade-Offs

Most households face trade-offs when shopping for groceries. They have grocery needs specific to their household size and brand preferences. They also may face a number of costs and constraints limiting their ability to fully optimize their store visits and purchases.

The budget constraint is one key component influencing household grocery shopping behavior, particularly for the lowest income consumers (Carlson et al. (2021)). To ensure their most important grocery needs are met within their budget, a lower income household may seek out stores that offer reliably low prices, either through lower base prices of particular products or through a lower priced assortment of products (i.e., smaller package sizes or lower-priced brands).

A related but distinct second set of constraints governs a household's ability to intertemporally shift their grocery expenditures to take advantage of temporary price discounts. Unless intertemporal price discounts are applied predictably and uniformly across product categories, a household's ability to avail themselves of intertemporally lower prices in stores that practice HiLo pricing ("HiLo stores") will depend on their travel costs, opportunity costs of time and the extent to which their grocery budget creates a binding constraint ("static budget constraint"). The lowest income households for whom the static budget constraint binds at the trip level may find HiLo stores unappealing if the average store price is sufficiently high: an already binding budget constraint would imply additional forgone necessities in no-discount weeks. Additionally, if a household has significant travel costs or high opportunity costs of time, they may not find it optimal to travel to a HiLo store sufficiently frequently to purchase their groceries at their lowest available prices. Although travel costs may be higher for lower income households (Allcott et al. (2019)), opportunity costs of time increase with income, suggesting HiLo stores also may not be particularly appealing for the highest income consumers. On the other hand, the highest income households are likely to be less price sensitive, making this the HiLo pricing strategy a less significant factor in their overall store choice and a more effective tool for price discrimination by retailers.

Finally, a third set of constraints governs a household's ability to take up quantity or volume discounts. Households with high storage costs may not be able to purchase the large package sizes required to obtain quantity discounts and therefore find club stores unappealing.

Moreover, households with grocery budget constraints that bind at the trip level may not be able to avail themselves of quantity discounts because doing so may mean leaving other grocery needs unmet. Both storage and budget constraints are more likely to bind for lower income households, therefore making stores that offer low prices by carrying large package sizes particularly unappealing to these households.

B Supplementary Data Description

Table B.1 shows that, depending on the year, our price indices are constructed using ~105-140 thousand unique UPCs at ~32-33 thousand stores and 85-123 unique retailers, covering 93-99% of the stores and 91-99% of the retailers in the data across 97-100% of the counties in the RMS data (Table B.2). Overall, we capture price indices for 8.7-9.7 million store-categories. Although these combinations represent only 42-45% of the unique store-categories observed in the data, they account for 87-91% of the total sales revenue (Table B.2). This suggests that the store-categories for which we do not capture price indices are relatively small and unlikely to be significant drivers of store or shopping choices for consumers.

Table B.1: Price Indices Data Summary: RMS Data

	Observations	Outlets		Geos	Products	
Year	Store-Categories	Stores	Retailers	Counties	Categories	UPCs
2012	9,190,154	32,748	101	2,476	865	104,516
2013	9,370,506	33,157	99	2,502	880	109,446
2014	9,275,589	32,911	96	2,511	873	111,990
2015	9,366,672	32,782	95	2,481	883	113,328
2016	9,211,759	32,965	97	2,535	871	115,212
2017	8,661,103	31,394	85	2,522	873	111,475
2018	9,744,827	33,233	123	2,556	881	135,340
2019	9,562,944	31,373	115	2,475	885	134,410
2020	9,606,227	31,635	118	2,481	948	139,596

This table provides an overview of the data used to construct the RMS Only price indices, focusing on the number of unique UPCs on which our price index calculations are based as well as the unique categories, retail outlets and geographies for which we recover price indices. Table B.2 complements this table by detailing the share of total revenue, as well as the proportion of unique categories, retail outlets, and geographic areas represented by these data.

Table B.2: Price Indices Data Summary: Share of RMS Data Captured

Year	Observations: Revenue	Observations: Store-Categories	Outlets: Stores	Outlets: Retailers	Geos: Counties	Products: Categories
2012	87%	42%	96%	94%	98%	91%
2013	87%	42%	96%	96%	99%	92%
2014	89%	42%	96%	93%	98%	91%
2015	91%	43%	97%	97%	99%	92%
2016	88%	43%	99%	99%	100%	91%
2017	89%	43%	99%	97%	100%	91%
2018	88%	44%	97%	91%	100%	91%
2019	87%	44%	93%	93%	97%	92%
2020	89%	45%	98%	96%	99%	85%

This table complements Table B.1 by detailing the share of total revenue, as well as the proportion of unique categories, retail outlets, geographic areas and store-categories represented by the recovered RMS Only store-category price indices.

C Estimation Details

C.1 Price Indices based on RMS and HMS data

We estimate the price indices in two steps. In the first step, we rely exclusively on the RMS data, which, as described in Section 2, contain the richer price panel. In the second step, we use the findings from step one in tandem with RMS and HMS data to estimate the pricing indices for all retailers in household's choice sets: those of participating retailers in the RMS data as well as those of non-participant retailers that only appear in the HMS data.

Step 1: Category-Store Pricing Strategies using RMS Data Although the RMS data constitute a rich price panel, in this data set, NielsenIQ only records a product price if the product was purchased at the store in the given week. We ensure that our price indices are based on shelf prices rather than prices conditional on purchase by first imputing the product prices on weeks in which they are missing and then estimating the price indices. We impute a missing price for a given product by taking the maximum across non-missing price observations for the same product at the same store within a moving 9-week time window. That is, in a given week in which a product price is missing, we assume that the shelf price would have been the maximum price of those observed in the preceding and following 4 weeks. This approach assumes that on weeks when the product is not purchased, the product price was likely a product's "base price," which is typically the higher of the observed prices in neighboring time periods (Hitsch et al. (2021)).[1]

We base our estimation of each of the four indices on only those products that are observed to be purchased at least 1,000 times in the RMS data. Furthermore, we only consider store-products where we observe prices in at least 39 out of the 52 weeks in a month, and only consider UPCs that are observed in at least 200 stores to ensure there is a valid comparison set. This selection criterion ensures a common point of reference in regressions 1 and 2 that control for product-specific price levels, and we apply it uniformly across the four measures to ensure measure comparability.

Step 2: State-Retailer Pricing Strategies using RMS and HMS Data The Step 1 estimation exercise yields a set of store-category-year specific price index estimates for the RMS-participant retailers, which we discuss in detail in Section 2. Although these retailers include many major industry players (Table B.1), they represent only a subset of the grocery retailers in a household's actual choice set – an important consideration for our store choice analysis. Hence, in estimation Step 2 we supplement RMS data with its HMS counterpart to calculate the price indices of both the RMS-participant retailers and retailers only observed in the HMS data.

The primary challenge in calculating price indices from HMS data is the sparsity of observations. Prices for a specific UPC are only recorded when a panelist purchases the item during a store visit. Our analysis in Section 3.3 provides a key insight that allows us to mitigate this sparsity: since most of the variation in pricing strategies is across retailer-state-years (Table D.3), we can reasonably estimate price indices by aggregating data across stores within each state. First, we construct a product-retailer-state-week price series by averaging observed prices within each retailer-state-week.[2] We then impute missing prices using the same method as in estimation Step 1, now including all retailer-product observations for which a price is observed in at least 5 weeks of the year.

In our Step 2 estimation, we derive the price indices of RMS-participant retailers from solely RMS data, as HMS data are already included. Thus, the imputed HMS data allow us to obtain price indices for the non-participant retailers in a household's choice set. Calculating price

[1]Our approach resembles closely the price imputation approach of Moshary et al. (2023)

[2]For stores in the HMS data without a recorded location, we use the panelist's location as a proxy for the store's location.

indices for both RMS and HMS retailers simultaneously increases the number of observations and improves the robustness of the indices.

C.2 Regression Weights

Relative Price To ensure that that the price index puts more weight on the relative prices of popular products, we weigh the regression in equation 1 by the total units of product j sold across all retailers and geographies in the given year, divided by the total number of price observations of product j.

Minimum Monthly Price To ensure that this pricing strategy index puts more weight on the prices of popular products, we weigh the regression in equation 2 by the total units of product j across all retailers and geographies in the given year, divided by the total number of monthly minimum price observations for product j.

Minimum Unit Price To ensure that the minimum is not affected by data outliers, in constructing this index, we exclude UPCs in the bottom decile of revenue at each retailer. Moreover, we do not weigh the regression in equation 3. The existence of at least some cheap product options in a given category, regardless of how in-demand, should afford an opportunity for an extremely budget-constrained household to purchase in that category. Therefore, weighing all product options equally should yield a more informative minimum unit price index.

Minimum Volume Price To ensure that this pricing strategy measure puts more weight on the prices of popular brands (e.g., due to higher quality), we weigh the regression in equation 4 by the total units of brand b across all retailers and geographies in the given year, divided by the number of price observations for that brand in the data.

D Supplementary Analysis of Price Index Variation

Table D.3: Adjusted R^2: Regression of RMS-only Price Indices on Fixed Effects (2012-2020)

		Retailer = Retailer Code				Retailer = Store Code			
	Fixed Effects	Relative Price (1)	Minimum Price (2)	Min Monthly Price (3)	Min Volume Price (4)	Relative Price (5)	Minimum Price (6)	Min Monthly Price (7)	Min Volume Price (8)
1	Category	0.07	0.17	0.08	0.14				
2	Format	0.11	0.30	0.10	0.11				
3	Format × Category	0.50	0.60	0.48	0.40				
4	Retailer	0.20	0.36	0.17	0.17	0.22	0.39	0.19	0.20
5	Retailer + Category	0.28	0.52	0.25	0.32	0.30	0.55	0.27	0.33
6	Retailer × Category	0.69	0.74	0.66	0.60	0.69	0.83	0.67	0.68
7	Retailer × Group	0.51	0.57	0.48	0.38	0.55	0.64	0.52	0.44
8	Retailer × Cat × State	0.72	0.78	0.70	0.65				
9	Retailer × Cat × County	0.72	0.80	0.70	0.67				
10	Retailer × Cat × County + Year	0.73	0.80	0.70	0.67	0.69	0.84	0.67	0.68
11	Retailer × Cat × Year	0.92	0.81	0.90	0.78				
12	Retailer × Cat × State ×Year	0.97	0.85	0.95	0.84				
13	Retailer × Cat × County × Year	0.98	0.87	0.96	0.86				

This table presents the adjusted R^2 from regressions of the price indices on varying levels of fixed effects, weighted by total deflated revenue. Columns (1)-(4) correspond to regressions where the Retailer fixed effect is defined by the Retailer Code, and columns (5)-(8) correspond to regressions where the Retailer fixed effect is defined by the Store Code. The price indices are estimated using RMS data only using step 1 in Web Appendix C.

E Supplementary Analysis of Costly Differentiation

E.1 Assortment Differentiation

We validate that the relationship between pricing strategies and assortment measure documented in Section 4.1 is not solely due to differences across retailer formats by estimating these regressions on Grocery retailer data only. Table E.4 shows that the relationships documented in Table **??** also hold within the Grocery format alone.

E.2 Geographic Differentiation

To support the retailer location analysis in Section 4.2, we provide further evidence of retailers with different pricing strategies locating near distinct consumer segments. Specifically, for each retailer, we calculate the mean and standard deviation of household income and size among its shoppers (recorded in the HMS data) and examine how these variables relate to retailer format and pricing strategies.

Table E.4: Pricing Strategies and Assortment (Grocery Retailers Only)

Model:	Rel Price (1)	Min Month Price (Res) (2)	Min Price (3)	Vol Price (4)	HiLo (5)
Variables: **Presence of Product Departments** (Retailer)					
Has Fresh Produce (Random Weight)	0.03***	-0.01	0.02	0.03**	0.25***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.09)
Has Fresh Produce (SKU)	-0.08*	-0.02**	-0.16***	-0.09	0.02
	(0.04)	(0.01)	(0.06)	(0.06)	(0.10)
Has Refrigerated (Milk, Meat)	-0.09**	0.00	-0.15***	-0.11**	-0.24***
	(0.04)	(0.01)	(0.06)	(0.04)	(0.08)
Has Deli	0.09**	-0.01	0.08	0.14***	0.31**
	(0.04)	(0.01)	(0.07)	(0.04)	(0.13)
Has Alcoholic Bev	0.09***	0.00	0.10***	0.15***	0.41***
	(0.03)	(0.01)	(0.04)	(0.03)	(0.08)
Has Frozen Foods	0.18***	0.02**	0.15***	0.25***	0.26***
	(0.04)	(0.01)	(0.03)	(0.05)	(0.06)
Variables: **Presence of Private Label** (Retailer)					
Has Private Label	-0.06	-0.01	-0.02	-0.11*	-0.21
	(0.06)	(0.01)	(0.08)	(0.06)	(0.24)
Variables: **Assortment Size** (Retailer-Department)					
Log(Categories)	-0.08*	0.01	-0.12***	-0.01	-0.07
	(0.04)	(0.01)	(0.04)	(0.03)	(0.15)
Log(Brands)	0.01	-0.02***	0.06***	-0.02	0.08
	(0.01)	(0.01)	(0.02)	(0.01)	(0.10)
Log(Sizes)	-0.08***	0.02***	-0.11***	-0.05***	-0.43***
	(0.02)	(0.01)	(0.02)	(0.02)	(0.13)
Median Name Brand Quality	0.11*	0.00	0.45***	0.20***	0.56
	(0.06)	(0.02)	(0.07)	(0.07)	(0.47)
Variables: **Package Size** (Retailer-Department)					
Share Top Quartile Sizes	-0.64***	0.21***	-0.18	-0.88***	-4.27***
	(0.11)	(0.06)	(0.27)	(0.13)	(0.78)
Share Bottom Quartile Sizes	-0.20*	0.13**	-0.28	-0.12	-3.62***
	(0.11)	(0.07)	(0.17)	(0.11)	(1.22)
Fixed-effects					
Dept-Year-State-Format	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	359,874	359,874	359,874	359,874	359,874
Adjusted R^2	0.57232	0.41242	0.52899	0.50913	0.47405
Within R^2	0.15819	0.04535	0.12790	0.16373	0.09823

Clustered (Retailer & Year-State-Channel) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

This table presents the results of regressions of each of the pricing strategies on the retailer- and retailer-product department assortment measures defined in Section 4.1, controlling for retailer format-state-product department fixed effects for Grocery retailers only. The HiLo pricing strategy measure is an indicator of whether the retailer-product department has an above-zero relative price and below-zero minimum monthly price, indicating HiLo pricing.

Table E.5 shows that warehouse clubs – and retailers with low volume prices more broadly – primarily serve high-income, large-household consumers, while retailers offering low unit prices attract smaller, lower-income households. Retailers with shallower discounts tend to cater to larger families, who may value the predictability of the stable grocery basket price. Additionally, similar to retailers with low volume prices, HiLo-pricing retailers primarily serve higher-income consumers, but with smaller household sizes. These shoppers may not require bulk purchases from big-box retailers but have the flexibility to take advantage of intertemporal promotions.

Table E.6 provides further evidence that grocery retailers competing more directly with particular retailer formats tend to target similar consumer segments. In this analysis, for each grocery retailer, we compute the share of revenue from the four other distinct retailer formats within the retailer's overall geographic footprint (i.e., all counties where the retailer operates). We then regress these competing format shares on the retailer's pricing strategies, controlling for the overall number of distinct competitors (retail chains). The retailer-specific geographic footprint also allows us to control for local market demographics using county fixed effects.

The results indicate that grocery retailers operating in areas with a higher prevalence of discount stores, similarly to this retailer format, tend to offer shallower monthly discounts. By contrast, retailers facing greater competition from dollar stores maintain lower relative, minimum, and volume prices. Notably, grocery retailers competing more heavily with warehouse clubs are more likely to adopt HiLo pricing, reinforcing our earlier finding in Table E.5 that HiLo pricing primarily appeals to a similarly higher-income but smaller-household consumers unable to take full advantage of the bulk buys of the warehouse clubs.

Table E.5: Pricing Strategies and Demographics Segmentation

	Indicator of Format					Pricing Strategies				
	Discount Store	Dollar Store	Drug Store	Ware-house Club	Grocery	Rel Price	Min Price	Min Month Price (Res)	Vol Price	Hilo
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables										
Log(Average HH Income)	-0.66	-0.34**	-0.01	1.45**	-0.43	0.45***	1.67***	-0.04	-0.22	0.81***
	(0.48)	(0.16)	(0.08)	(0.58)	(0.59)	(0.11)	(0.43)	(0.03)	(0.24)	(0.21)
Log(StDev HH Income)	0.13	0.00	0.00	-0.20***	0.07	-0.05***	-0.17***	0.00	0.05***	-0.07**
	(0.08)	(0.01)	(0.01)	(0.04)	(0.07)	(0.01)	(0.03)	(0.00)	(0.02)	(0.03)
Log(Average HH Size)	0.58*	0.04	-0.37*	0.98***	-1.23***	-0.04	0.69**	0.06***	-0.58***	-0.61**
	(0.32)	(0.03)	(0.20)	(0.37)	(0.29)	(0.09)	(0.32)	(0.02)	(0.10)	(0.25)
Log(StDev HH Size)	0.22	-0.04**	-0.01	-0.37***	0.19	-0.06**	-0.24**	0.00	0.14***	0.01
	(0.16)	(0.02)	(0.02)	(0.12)	(0.17)	(0.03)	(0.10)	(0.01)	(0.04)	(0.08)
Fixed-effects										
Year-County	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics										
Observations	162,533	162,533	162,533	162,533	162,533	161,722	161,722	161,722	161,722	161,722
R^2	0.14206	0.06050	0.03093	0.21619	0.09238	0.18557	0.24046	0.15134	0.12027	0.20496

Clustered (Retailer & Year-County) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table E.6: Pricing Strategies and Competitor Formats

	Pricing Strategies (Grocery Only)				
	Rel Price	Min Price	Min Month Price (Res)	Vol Price	Hilo
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Discount Store Comps (Share of Sales)	0.03	-0.05	0.10***	0.15	-0.69
	(0.09)	(0.11)	(0.03)	(0.11)	(0.72)
Dollar Store Comps (Share of Sales)	-0.67**	-1.51***	0.17	-1.04***	-5.79*
	(0.34)	(0.44)	(0.13)	(0.38)	(2.96)
Drug Store Comps (Share of Sales)	0.06	0.14	0.04	0.17	0.00
	(0.14)	(0.19)	(0.04)	(0.17)	(1.07)
Warehouse Club Comps (Share of Sales)	0.42	0.56**	-0.26***	0.25	6.01**
	(0.26)	(0.27)	(0.09)	(0.25)	(2.81)
Number of Competitors	0.00	0.00	0.00***	0.00	-0.04**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
Fixed-effects					
Year-County	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	74,889	74,889	74,889	74,889	74,889
Adjusted R^2	0.47557	0.40658	0.42760	0.33304	0.45065

Clustered (Retailer & Year-County) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

F Retailer Visits Robustness & Alternative Specifications

F.1 Assortment Results

Table F.7: Store choice - Assortment Coefficients

	log(E(Num Weeks Visited Per Month))
Model:	(1)
Variables	
Income ($10,000)	0.002 (0.006)
Household Size	0.14*** (0.02)
Single Family Home	0.03 (0.06)
Is Household Head Not Working	0.51*** (0.04)
log(Num Brands)	0.68*** (0.008)
log(Num UPCs)	-0.002 (0.007)
log(Num Sizes)	-0.29*** (0.01)
Brand Quality	-0.28*** (0.02)
% UPC in Top Quartile Sizes	1.4*** (0.04)
% UPC in Bot Quartile Sizes	2.0*** (0.05)
Income ($10,000) × Random Weight Produce Avail.	-0.005*** (0.0006)
Income ($10,000) × Fresh Produce Avail.	0.01*** (0.0007)
Income ($10,000) × Refrigerated Avail.	-0.003* (0.002)
Income ($10,000) × Deli Avail.	0.0003 (0.0008)
Income ($10,000) × Alcohol Avail.	0.007*** (0.0005)
Income ($10,000) × Frozen Foods Avail.	-0.004*** (0.0009)
Income ($10,000) × Private Label Avail.	0.01** (0.005)
Income ($10,000) × log(Num Brands)	-0.001* (0.0007)
Income ($10,000) × log(Num UPCs)	-0.001** (0.0006)
Income ($10,000) × log(Num Sizes)	-0.007*** (0.0008)
Income ($10,000) × Brand Quality	0.06*** (0.002)
Income ($10,000) × % UPC in Top Quartile Sizes	-0.0009 (0.002)
Income ($10,000) × % UPC in Bot Quartile Sizes	0.02*** (0.004)
Household Size × Random Weight Produce Avail.	-0.007*** (0.002)
Household Size × Fresh Produce Avail.	0.01*** (0.002)
Household Size × Refrigerated Avail.	0.04*** (0.004)
Household Size × Deli Avail.	0.03*** (0.002)
Household Size × Alcohol Avail.	-0.02*** (0.002)
Household Size × Frozen Foods Avail.	-0.002 (0.003)
Household Size × Private Label Avail.	-0.08*** (0.02)
Household Size × log(Num Brands)	-0.01*** (0.002)
Household Size × log(Num UPCs)	-0.04*** (0.002)
Household Size × log(Num Sizes)	0.05*** (0.002)
Household Size × Brand Quality	-0.002 (0.005)
Household Size × % UPC in Top Quartile Sizes	0.06*** (0.007)
Household Size × % UPC in Bot Quartile Sizes	0.02* (0.01)
Single Family Home × Random Weight Produce Avail.	0.04*** (0.005)
Single Family Home × Fresh Produce Avail.	0.01** (0.006)
Single Family Home × Refrigerated Avail.	0.03** (0.01)
Single Family Home × Deli Avail.	0.06*** (0.007)
Single Family Home × Alcohol Avail.	0.02*** (0.005)
Single Family Home × Frozen Foods Avail.	-0.03*** (0.008)
Single Family Home × Private Label Avail.	-0.21*** (0.05)
Single Family Home × log(Num Brands)	-0.01** (0.006)
Single Family Home × log(Num UPCs)	0.06*** (0.006)
Single Family Home × log(Num Sizes)	-0.08*** (0.007)
Single Family Home × Brand Quality	0.14*** (0.01)
Single Family Home × % UPC in Top Quartile Sizes	0.32*** (0.02)
Single Family Home × % UPC in Bot Quartile Sizes	-0.10*** (0.03)
Is Household Head Not Working × Random Weight Produce Avail.	0.10*** (0.004)
Is Household Head Not Working × Fresh Produce Avail.	-0.07*** (0.005)
Is Household Head Not Working × Refrigerated Avail.	-0.08*** (0.01)
Is Household Head Not Working × Deli Avail.	-0.07*** (0.005)
Is Household Head Not Working × Alcohol Avail.	0.01*** (0.004)
Is Household Head Not Working × Frozen Foods Avail.	-0.08*** (0.006)
Is Household Head Not Working × Private Label Avail.	0.06 (0.04)
Is Household Head Not Working × log(Num Brands)	-0.02*** (0.005)
Is Household Head Not Working × log(Num UPCs)	-0.02*** (0.005)
Is Household Head Not Working × log(Num Sizes)	-0.02*** (0.006)
Is Household Head Not Working × Brand Quality	-0.30*** (0.01)
Is Household Head Not Working × % UPC in Top Quartile Sizes	0.25*** (0.02)
Is Household Head Not Working × % UPC in Bot Quartile Sizes	-0.06** (0.03)
Fixed-effects	
Year-Month-ZIP Code-Retailer	Yes
Fit statistics	
Standard-Errors	Year-Month-ZIP Code-Retailer
Observations	26,195,119
Pseudo R^2	-0.02794

Clustered (Year-Month-ZIP Code-Retailer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table F.8: Store choice - Assortment Coefficients

Model:	(1)	(2)	(3)
Variables			
Income ($10,000)	0.007*** (0.0001)	-0.0001 (0.006)	-0.004 (0.006)
Household Size	0.010*** (0.0004)	0.13*** (0.02)	0.14*** (0.02)
log(Num Brands)		0.65*** (0.007)	0.67*** (0.007)
log(Num UPCs)		0.07*** (0.006)	0.04*** (0.006)
log(Num Sizes)		-0.39*** (0.009)	-0.38*** (0.009)
Brand Quality		-0.63*** (0.01)	-0.40*** (0.02)
% UPC in Top Quartile Sizes		1.7*** (0.03)	1.8*** (0.04)
% UPC in Bot Quartile Sizes		2.5*** (0.04)	2.0*** (0.04)
Income ($10,000) × Random Weight Produce Avail.		-0.004*** (0.0006)	-0.007*** (0.0006)
Income ($10,000) × Fresh Produce Avail.		0.005*** (0.0007)	0.01*** (0.0007)
Income ($10,000) × Refrigerated Avail.		0.0004 (0.002)	-0.0003 (0.002)
Income ($10,000) × Deli Avail.		-0.006*** (0.0007)	0.003*** (0.0007)
Income ($10,000) × Alcohol Avail.		0.01*** (0.0005)	0.007*** (0.0005)
Income ($10,000) × Frozen Foods Avail.		-0.0006 (0.0008)	-0.003*** (0.0008)
Income ($10,000) × Private Label Avail.		0.004 (0.005)	0.005 (0.005)
Income ($10,000) × log(Num Brands)		0.005*** (0.0006)	-0.0009 (0.0007)
Income ($10,000) × log(Num UPCs)		-0.0002 (0.0006)	-0.0006 (0.0006)
Income ($10,000) × log(Num Sizes)		-0.01*** (0.0007)	-0.009*** (0.0007)
Income ($10,000) × Brand Quality		0.12*** (0.002)	0.07*** (0.002)
Income ($10,000) × % UPC in Top Quartile Sizes		0.05*** (0.002)	-0.001 (0.002)
Income ($10,000) × % UPC in Bot Quartile Sizes		-0.04*** (0.004)	0.02*** (0.004)
Household Size × Random Weight Produce Avail.		-0.01*** (0.002)	-0.006*** (0.002)
Household Size × Fresh Produce Avail.		0.04*** (0.002)	0.02*** (0.002)
Household Size × Refrigerated Avail.		0.03*** (0.004)	0.04*** (0.004)
Household Size × Deli Avail.		0.03*** (0.002)	0.03*** (0.002)
Household Size × Alcohol Avail.		-0.03*** (0.002)	-0.02*** (0.002)
Household Size × Frozen Foods Avail.		0.005** (0.003)	0.0003 (0.003)
Household Size × Private Label Avail.		-0.11*** (0.02)	-0.09*** (0.02)
Household Size × log(Num Brands)		-0.03*** (0.002)	-0.009*** (0.002)
Household Size × log(Num UPCs)		-0.04*** (0.002)	-0.04*** (0.002)
Household Size × log(Num Sizes)		0.06*** (0.002)	0.04*** (0.002)
Household Size × Brand Quality		0.02*** (0.004)	0.01** (0.005)
Household Size × % UPC in Top Quartile Sizes		0.13*** (0.004)	0.06*** (0.007)
Household Size × % UPC in Bot Quartile Sizes		-0.02* (0.01)	0.02* (0.01)
Assortment Controls	No	Yes	Yes
Assortment-Income Interactions	No	Yes	Yes
thisHeader	log(E(Num Weeks Visited Per Month))	log(E(Num Weeks Visited Per Month))	log(E(Num Weeks Visited Per Month))
Fixed-effects			
Year-Month-ZIP Code-Retailer	Yes	Yes	Yes
Fit statistics			
Standard-Errors		Year-Month-ZIP Code-Retailer	
Observations	26,448,961	26,195,119	26,195,119
Pseudo R^2	-0.03216	-0.03013	-0.02956

Clustered (Year-Month-ZIP Code-Retailer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

F.2 Retailer Visits and Other Household Characteristics

We expand our descriptive model of retailer visits by incorporating two additional household characteristics: whether there is a household head not in the labor force and whether they residence in a single-family home. We hypothesize that households where one adult is not working will be better able to take advantage of price promotions due to more time being available for shopping, while those in single-family homes will be better positioned to take advantage of larger package sizes and price promotions due to greater storage capacity.

Table F.9 presents these results. The relationships between price indices, income, and household size remain qualitatively similar to our previous findings. Households where a head is not in the labor force prefer price variation and lower unit prices, perhaps reflecting that they can more easily engage in smaller stock up trips. Households in single-family homes prefer both price variation and lower volume prices.

To illustrate these effects, we examine how different price metrics influence various household

Table F.9: Retailer Visits, Pricing Strategies, and Household Characteristics (Addl)

Model:	Weeks Visited in Month (1)
Variables	
Income ($10,000)	0.010
	(0.006)
Household Size	0.12***
	(0.02)
Household Head Not in Labor Force	0.44***
	(0.04)
Single Family Home	0.08
	(0.06)
Relative Price	0.65***
	(0.08)
Min Monthly Price	-0.51***
	(0.07)
Min Unit Price	0.07***
	(0.01)
Min Vol Price	-0.008
	(0.02)
Income ($10,000) × Relative Price	0.15***
	(0.004)
Income ($10,000) × Min Monthly Price	-0.17***
	(0.004)
Income ($10,000) × Min Unit Price	0.08***
	(0.001)
Income ($10,000) × Min Vol Price	0.02***
	(0.002)
Household Size × Relative Price	-0.52***
	(0.01)
Household Size × Min Monthly Price	0.42***
	(0.01)
Household Size × Min Unit Price	0.03***
	(0.003)
Household Size × Min Vol Price	-0.02***
	(0.005)
Household Head Not in Labor Force × Relative Price	0.59***
	(0.03)
Household Head Not in Labor Force × Min Monthly Price	-0.20***
	(0.03)
Household Head Not in Labor Force × Min Unit Price	-0.27***
	(0.007)
Household Head Not in Labor Force × Min Vol Price	0.002
	(0.01)
Single Family Home × Relative Price	0.72***
	(0.04)
Single Family Home × Min Monthly Price	-0.68***
	(0.04)
Single Family Home × Min Unit Price	0.01*
	(0.009)
Single Family Home × Min Vol Price	-0.19***
	(0.01)
Fixed-effects	
Year-Month-ZIP Code-Retailer	Yes
Fit statistics	
Observations	26,291,862
Squared Correlation	0.28844
Pseudo R^2	-0.02775
BIC	9.7×10^{10}

Clustered (Year-Month-ZIP Code-Retailer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

This table presents the results of Poisson regressions analyzing the relationship between household visit frequency to a retailer per month and retailer pricing strategies, along with consumer demographics. The full specification in Column 3 follows equation 7 and includes an expanded set of household characteristics beyond those in Table 6.

types. When holding other pricing metrics constant, increased price variation primarily attracts small, high-income households in single-family homes with a head of household outside the labor force. For instance, a household with \$100,000 annual income, two members, a stay-at-home parent, and single-family residence will respond to a 10% decrease in the minimum monthly price by increasing their visits by 22.5%. In contrast, a household with \$20,000 annual income, four members, no stay-at-home parent, and apartment residence will decrease their visits by 8.3%.

For most consumers in our sample, lower unit prices prove unattractive when other pricing factors remain constant, as smaller package sizes are generally impractical. However, single-person households with low income (\$10,000) and a head not in the workforce show a slight increase in visit frequency in response to reduced minimum unit prices.

Lower volume prices primarily appeal to households with lower incomes, larger sizes, and single-family residences. A household with \$20,000 annual income, six members, a single-family home, and dual working parents will increase store visits by 2.78% in response to a 10% decrease in volume prices. Conversely, a household with \$100,000 annual income, two members, apartment residence, and dual working parents would reduce visits by 1.54% under the same price change. These heterogeneous responses to changes highlight that consumers put differing value on different pricing metrics in a manner consisent with the constraints their demographic group may face.

F.3 Attenuation Bias

A potential limitation of our retailer visit model is that the price indices were estimated using a limited dataset, potentially introducing measurement error, particularly for retailers appearing exclusively in the HMS data. To address this concern, we first re-estimate our model using only retailers present in the RMS data, which provides a more comprehensive panel of retail pricing. While we anticipate some coefficient changes due to the modified retailer sample, we expect the qualitative relationships between consumer demographics and price indicies to remain consistent. Second, we re-estimate our Retailer visit model focusing only on zip codes where at least 10 household panelists are present. This number of households would result in additional price observations for the set of available retailers in the HMS.

Table F.10: Store Choice on RMS Only

Model:	Num Weeks Visited Per Month (1)	(2)
Variables		
Income ($10,000)	0.03**	0.03***
	(0.01)	(0.008)
Relative Price	4.7***	1.8***
	(0.19)	(0.10)
Min Monthly Price	-5.2***	-1.4***
	(0.19)	(0.09)
Min Unit Price	0.64***	-0.07***
	(0.04)	(0.02)
Min Vol Price	0.26***	-0.09***
	(0.02)	(0.02)
Income ($10,000) × Relative Price	0.006	0.14***
	(0.009)	(0.005)
Income ($10,000) × Min Monthly Price	-0.06***	-0.16***
	(0.01)	(0.005)
Income ($10,000) × Min Unit Price	0.14***	0.06***
	(0.002)	(0.002)
Income ($10,000) × Min Vol Price	-0.02***	0.03***
	(0.002)	(0.002)
Relative Price × Household Size	-0.62***	-0.62***
	(0.03)	(0.01)
Min Monthly Price × Household Size	0.40***	0.54***
	(0.03)	(0.01)
Min Unit Price × Household Size	0.09***	0.04***
	(0.003)	(0.002)
Min Vol Price × Household Size	-0.01**	-0.05***
	(0.005)	(0.005)
Assortment Controls	No	Yes
Assortment-Income Interactions	No	Yes
RMS Data Only	Yes	No
At least 10 households in Zip	No	Yes
Fixed-effects		
Year-Month-ZIP Code-Retailer	Yes	Yes
Fit statistics		
Observations	9,324,436	14,546,886
Pseudo R^2	-0.03600	-0.01917

Clustered (Year-Month-ZIP Code-Retailer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table F.10 presents these results. Notably, all significant interactions between consumer demographics and price indices maintain their original signs, with one exception: the interaction between income and volume price now exhibits a negative coefficient in the first analysis. Therefore we do not believe measurement error in the price indices can bias our results.

F.4 Transportation

Transportation access represents a significant constraint that may influence household grocery shopping behavior. Households with car access may be more inclined to patronize retailers offering larger package sizes and better volume prices. Since transportation access potentially correlates with income and household size, it could affect the estimated relationships between these demographic factors, our pricing metrics, and store visit decisions.

Table F.11: Store Choice with Zip Code Splits

	Vehicle Availability		Work Travel Time		Transit, Bike or Walk to Work	
	Below	Above	Below	Above	None	Some
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Income ($10,000)	-0.01*	-0.04***	0.10**	-0.03***	-0.03	-0.03***
	(0.007)	(0.010)	(0.04)	(0.006)	(0.02)	(0.006)
Household Size	0.08***	0.23***	0.14	0.12***	0.04	0.13***
	(0.02)	(0.03)	(0.28)	(0.02)	(0.06)	(0.02)
Relative Price	2.0***	-0.04	1.9***	1.4***	-0.65***	1.4***
	(0.09)	(0.13)	(0.69)	(0.08)	(0.21)	(0.08)
Min Monthly Price	-1.6***	0.49***	-1.7***	-1.1***	1.7***	-1.1***
	(0.08)	(0.12)	(0.65)	(0.07)	(0.20)	(0.07)
Min Unit Price	0.07***	0.01	-0.42***	0.05***	-0.15***	0.03**
	(0.02)	(0.02)	(0.11)	(0.01)	(0.04)	(0.01)
Min Vol Price	0.03	-0.15***	0.07	-0.009	-0.25***	0.005
	(0.02)	(0.03)	(0.17)	(0.02)	(0.05)	(0.02)
Income ($10,000) × Relative Price	0.13***	0.17***	0.23***	0.14***	0.21***	0.14***
	(0.005)	(0.006)	(0.04)	(0.004)	(0.01)	(0.004)
Income ($10,000) × Min Monthly Price	-0.16***	-0.18***	-0.15***	-0.17***	-0.24***	-0.17***
	(0.005)	(0.006)	(0.04)	(0.004)	(0.01)	(0.004)
Income ($10,000) × Min Unit Price	0.06***	0.05***	0.09***	0.06***	0.03***	0.06***
	(0.001)	(0.002)	(0.01)	(0.001)	(0.003)	(0.001)
Income ($10,000) × Min Vol Price	0.01***	0.03***	-0.06***	0.02***	0.02***	0.01***
	(0.002)	(0.003)	(0.02)	(0.002)	(0.005)	(0.002)
Household Size × Relative Price	-0.49***	-0.57***	-0.88***	-0.52***	-0.47***	-0.53***
	(0.01)	(0.02)	(0.13)	(0.01)	(0.04)	(0.01)
Household Size × Min Monthly Price	0.39***	0.42***	0.76***	0.42***	0.19***	0.43***
	(0.01)	(0.02)	(0.12)	(0.01)	(0.03)	(0.01)
Household Size × Min Unit Price	0.02***	0.04***	-0.15***	0.02***	0.09***	0.02***
	(0.004)	(0.006)	(0.03)	(0.004)	(0.01)	(0.004)
Household Size × Min Vol Price	-0.04***	-0.07***	0.07	-0.06***	0.02	-0.06***
	(0.007)	(0.009)	(0.05)	(0.006)	(0.02)	(0.006)
Fixed-effects						
Year-Month-ZIP Code-Retailer	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	15,200,707	9,513,426	684,016	20,788,968	3,168,248	20,294,274
Pseudo R^2	-0.03705	-0.03732	-0.56370	-0.03395	-0.15023	-0.03335

Clustered (Year-Month-ZIP Code-Retailer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Lacking direct data on household transportation options, we examined how shopping decisions vary according to the transportation characteristics of the zip code the household resides in. We extracted three metrics from the 2015 and 2020 US Census: per capita vehicle availability, average work commute time, and per capita use of alternative transportation (public transit, biking, and walking to work). We then divided our sample into zip codes above and below the median for each characteristic and replicated our analysis for each subsample.[3]

Table F.11 presents these results. The interactions between our pricing metrics and both income and household size remain qualitatively consistent across almost all specifications. Moreover, the interaction coefficients are similar for households in zip codes both above and below the median for each transportation metric. This consistency strengthens our confidence that the relationships identified in our initial analysis directly reflect the influence of income and household size, rather than stemming from omitted transportation variables.

F.5 Discrete Choice Model

In this section, we develop a structural model of consumer store choice using our pricing metrics, estimated through logit inversion and linear regression. This approach offers a significant advantage over our descriptive model by better capturing substitution patterns between competing grocery retailers. However, this structural approach introduces several limitations.

First, unlike our descriptive model, it cannot accommodate a rich set of fixed effects necessary to account for quality differences across retailers and the specific assortment of retailers available within each zip code. Second, it requires accurately estimated choice shares and can only incorporate one demographic variable at a time to ensure sufficient data in each demographic segment. Third, it necessitates a more aggregated approach to pricing metrics, which are consolidated to the state-retailer level using the aggregate departmental weightings from the full dataset, rather than reflecting household-specific consumption preferences.

We first specify consumer utility as follows:

$$u_{iry} = \beta_I I_{iy} + (\beta^{\mathbf{P}} + \beta^{\mathbf{P}}_{\mathbf{I}} I_{iy})(\mathbf{p}^{\text{rel}}_{rsy} + \mathbf{p}^{\text{mon}}_{rsy} + \mathbf{p}^{\text{vol}}_{rsy} + \mathbf{p}^{\text{unit}}_{rsy}) + \gamma_{sy} + \gamma_{ry} + \xi_{rsy} + \varepsilon_{isy} \tag{F.1}$$

where u_{iry} represents the utility the consumer i gets from visiting retailer r in year y, I_{iy}

[3]The resulting subsamples have unequal sizes due to population density variations across zip codes.

Table F.12: Store Visits in Discrete Choice Model

Dependent Variable:	Visit Shares
Model:	(1)
Variables	
Relative Price	-2.6***
	(0.50)
Min Monthly Price	2.2***
	(0.48)
Min Unit Price	-0.55***
	(0.09)
Min Vol Price	0.43***
	(0.13)
Income Quartile	-0.48***
	(0.009)
Relative Price × Income Quartile	0.31***
	(0.08)
Min Monthly Price × Income Quartile	-0.18**
	(0.08)
Min Unit Price × Income Quartile	0.14***
	(0.01)
Min Vol Price × Income Quartile	-0.18***
	(0.02)
Fixed-effects	
State-Year	Yes
Retailer-Year	Yes
Fit statistics	
Observations	100,912
R^2	0.70660
Within R^2	0.40931

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

represents the income quartile of consumer i in year y, γ_{sy} and γ_{ry} represent fixed effects at the state-year and retailer-year level, ξ_{rsy} is an idiosyncratic mean-0 shock to consumer utilities a the retailer-state-year level, an IID ε_{isy} has a type-1 extreme value distribution. The outside good represents not visiting a retailer, and its utility is normalized to 0. We specify the market size as 52 total visits in a year per household.

We replace the price indices with their estimated counterparts and use a log-share inversion to develop our estimating equation:

$$log(s_{Iry}) - log(s_{I0y}) = \beta_I I_{Iy} + (\beta^{\mathbf{p}} + \beta^{\mathbf{p}}_{\mathbf{I}} I_{Iy})(\hat{\mathbf{p}}^{\text{rel}}_{rsy} + \hat{\mathbf{p}}^{\text{mon}}_{rsy} + \hat{\mathbf{p}}^{\text{vol}}_{rsy} + \hat{\mathbf{p}}^{\text{unit}}_{rsy}) + \gamma_{sy} + \gamma_{ry} + \xi_{rsy} \quad \text{(F.2)}$$

where s_{Iry} is the market share of retailer r among households with income I during year y, and s_{I0y} represents the share of the outside good. To ensure the choice shares are estimated accurately, we only use state-income cells with more than 30 observations.

We estimate this equation using linear regression, with results presented in Table F.12. Our findings indicate that households across all income levels avoid retailers with high relative prices, with this aversion most pronounced among lower-income consumers. Higher-income consumers demonstrate a relative preference for retailers offering greater price variation and lower volume prices. Conversely, lower-income consumers exhibit stronger preferences for retailers with lower unit prices. These patterns align qualitatively with the findings from our descriptive model, providing consistent evidence across methodological approaches.

G Basket Savings Calculation Details

First, we exclude likely "filler trips," defined as any shopping basket with a total expenditure of less than \$50.[4] We then focus our attention to shopping baskets for which we have a measure of the price at the given retailer-county-week in the NielsenIQ RMS data for each product in the basket. We impute shelf prices on no-purchase weeks as in Section 3. Additionally, to ensure a reliable counterfactual comparison set, we exclude from the shelf price panel any product that has non-zero sales in less than 26 out of 52 weeks in that year, which limits the number of prices matched to the realized purchase baskets, as described above.

We use the NielsenIQ RMS price, instead of the price recorded in the trips data, as the measure of the paid price to allow for a direct comparison between paid and counterfactual prices. Because the counterfactual prices are also drawn from the same NielsenIQ RMS price panel, using NielsenIQ RMS-only prices ensures that the calculation of foregone savings is unaffected by discrepancies in how prices are recorded between the two datasets.

Matching shelf prices for private label products from NielsenIQ RMS to the NielsenIQ trips data data poses a challenge, as the product identifier (UPC) for private label items differs between the RMS and HMS datasets. Therefore, for private label products, we match prices to the items in the basket based on retailer-specific private label brand-size combinations rather than using UPCs.

In columns (1)-(3) of Table G.13, we show the numbers of unique households, retailers visited and shopping trips by income quintile in the resulting data set. In columns (4)-(5), we report the average number of monthly trips and the average monthly grocery expenditures for in-sample households across all in-sample categories and store formats, with the unique baskets

[4]We arrive at this approximate number by dividing the average annual food budget of the lowest income quintile households (\$5,090 in 2022, Martin (2024)) by their average number of grocery shopping trips in a year (12*8.6, Table G.13).

Table G.13: Basket Savings Analysis Data Summary

HH Income	Unique HH (1)	Unique Retailers (2)	Unique Baskets (3)	Avg. Monthly Baskets (4)	Avg. Monthly Spend (5)
≤25,000	24,735	150	710,480	8.6	243.8
$25,000-50,000	42,739	152	1,436,816	8.6	271.6
$50,000-80,000	48,895	152	1,859,679	8.3	296.7
$80,000-130,000	37,038	151	1,499,630	8.2	318.3
$130,000+	16,939	106	675,365	7.9	323.8

Columns 1-3 of this table show the unique households, retailers and grocery baskets by household income used in the foregone savings analysis. Columns 4-5 report the average number of monthly trips and the average monthly grocery expenditures for in-sample households across all in-sample categories and store formats, with the unique baskets from column (3) representing a subset of such trips. Income quintiles are formed as described in Section 2.

Table G.14: Foregone Savings Potential by Household Income

			Within Retailer Foregone Savings					Across Retailer Foregone Savings	
			Inter-temporal		Brand		Volume	Retailer	
HH Income	Avg. Monthly Baskets	Avg. Monthly Spend	Savings % (1)	Addl Weeks (2)	Savings % (3)	% to Private (4)	Savings % (5)	Savings % (6)	Addl Retailers (7)
≤25,000	8.6	243.8	10.8	1.8	20.1	16.7	9.3	6.9	1.9
$25,000-50,000	8.6	271.6	10.8	1.9	20.6	17.1	9.4	6.8	1.9
$50,000-80,000	8.3	296.7	10.5	1.9	21.1	17.4	9.4	7.0	2.1
$80,000-130,000	8.2	318.3	10.1	2.0	21.9	18.0	9.5	7.6	2.2
$130,000+	7.9	323.8	9.9	2.0	21.6	18.9	9.1	7.9	2.2

This table summarizes the percentage inter-temporal, brand, volume and retailer savings foregone by households of different incomes and the corresponding additional grocery basket changes effecting such savings would require. Income quintiles are formed as described in Section 2.

from column (3) representing a subset of such trips.