Income and Retail Prices within a City: Evidence on Product Composition and Market Segmentation*

Matías Brum†, Iael Klaczko‡ and Leandro Zipitría§

This version: May 8, 2026

Abstract

This paper studies how retail prices vary with neighborhood income within a city. Using detailed store-product data for Montevideo, we document a positive income-price relationship in aggregate comparisons. However, this relationship declines sharply as comparisons are restricted to more homogeneous goods and retail environments, and becomes economically negligible for identical products sold by the same retailer. This pattern reveals that the observed income-price elasticity reflects differences in product composition and retail structure across neighborhoods, rather than systematic price differences for identical goods. The remaining non-linearity is concentrated at the bottom of the income distribution and is driven primarily by chain retailers. Differences in local competition and product assortment account for part of this heterogeneity but do not eliminate it. Overall, the results indicate that spatial price inequality within cities operates

*We thank the Consumer Defense Unit of the Uruguayan Ministry of Economy and Finance—especially Alejandro Mateos and Gabriela Carrasco—for granting access to the supermarket price database, and Scanntech—especially Martín Severgnini—for providing the grocery price database. We are also grateful to Guillermo Vera for helping us map the stores to neighborhoods. We thank Fernando Borraz and Martín Rama for helpful comments and suggestions, and participants at *Jornadas Anuales de Economía, Banco Central del Uruguay*, the 2024 LACEA-LAMES, and the SEU meetings. All remaining errors remain ours.

†Universidad ORT Uruguay.

‡Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República. iael.klaczko@cienciassociales.edu.uy

§Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República. leandro.zipitria@cienciassociales.edu.uy

mainly through the composition of available goods and retail environments, not through differential pricing of comparable products.

JEL CODES: D31, D63, L11, L81, R21.

Keywords: Price Dispersion, Market Segmentation, Inequality.

1 Introduction

Do low-income neighborhoods face higher retail prices? A large literature has debated this question, reaching mixed conclusions depending on what is being compared. We show that the answer depends critically on the level of aggregation: when prices are compared broadly across neighborhoods, higher-income areas appear to have higher prices. But this relationship shrinks sharply as the comparison is restricted to more comparable goods and retail environments, and becomes economically negligible once identical products sold by the same retailer are compared. The implication is direct: spatial price inequality within cities is not primarily a story of retailers charging different prices for the same good depending on neighborhood income. It is, instead, a story of differences in which products are available and which types of stores operate across neighborhoods.

A key element in this literature is the presence of spatial frictions. Consumers do not freely arbitrage prices across locations; instead, they tend to shop close to their place of residence. Evidence from Eizenberg, Lach, and Oren-Yiftach (2021) shows that a substantial share of grocery expenditures occurs within the home neighborhood, even when prices are higher, highlighting the importance of distance and travel costs in shaping shopping behavior. Consistent with this evidence, a recent report for Uruguay shows that most consumers travel short distances to shop: in the full sample, the majority walk to the store, travel less than three blocks, and spend only three to five minutes getting there. The same report also shows that shopping trips predominantly start at home: 77% of consumers report traveling from home, 15% from work, and 8% from another location (Comisión de Promoción y Defensa de la Competencia, 2022). While larger purchases involve longer trips, everyday shopping remains highly localized, particularly for small and medium-sized stores.

These patterns imply that retail competition is geographically segmented at a very local level. In this context, the law of one price does not hold across neighborhoods, even within the same city, because consumers do not systematically reallocate their purchases in response to spatial price differences. As a result, equilibrium prices can differ persistently across neighborhoods.

This has a direct implication for our empirical approach. If markets are local, neighborhood income becomes a relevant determinant of the local demand environment retailers face. In turn, this can generate systematic differences in prices across neighborhoods with different income levels. We therefore interpret the income–price relationship as capturing how local market conditions, proxied by neighborhood income, are reflected in equilibrium pricing. Under this interpretation, estimating an income–price elasticity across neighborhoods is a meaningful way to characterize spatial price differences, rather than a reduced-form correlation that would be arbitraged away in an integrated market.

However, the answer depends critically on what is being compared. Differences in observed prices across locations may reflect variation in the prices of identical goods, as well as differences in product composition, store types, and local market conditions. Distinguishing between these channels requires detailed data at the product and store level.

In this paper, we study the relationship between neighborhood income and retail prices using detailed data for Montevideo, the capital and largest city of Uruguay. Montevideo provides a particularly useful setting for this analysis for three reasons. First, the city exhibits substantial variation in neighborhood income within a compact and relatively integrated urban market, allowing us to study spatial price differences while abstracting from cross-city variation in costs, wages, and supply conditions. Second, Uruguay offers unusually rich administrative data: our analysis combines two complementary price datasets covering both large supermarket chains and independent retailers, matched to neighborhood-level income from the national statistical office, yielding over three million store-product-month observations across 49 neighborhoods and nearly a decade. Third, retail markets in Montevideo are geographically segmented in a way that makes neighborhood income a meaningful determinant of the local demand environment. Survey evidence shows that most consumers shop within a few blocks of their homes, predominantly by walking, which implies that local retail conditions are not arbitraged away by consumer mobility. This combination of income variation, data richness, and spatial frictions makes Montevideo an informative laboratory for studying how retail prices vary across neighborhoods with different income levels.

Our empirical strategy exploits variation at different levels of aggregation. We first compare prices across neighborhoods without restrictions, then progressively narrow the comparison to within-category, within-product, and within-store-product comparisons. This approach allows us to distinguish between composition effects and differences in the prices of identical goods. While this approach is related to the decomposition between composition effects and price differences emphasized in the literature, our setting differs in two important dimensions. First, we study these margins within a single city, abstracting from cross-city cost differences and focusing on local retail environments. Second, we combine this decomposition with variation across retail formats, allowing us to assess how retailers' organizational features shape the income-price relationship.

We find a positive relationship between income and prices in aggregate comparisons. However, this relationship declines sharply as the comparison is restricted to more comparable goods and retail environments. This pattern is central to our interpretation of the results below.

We also document that the income-price relationship is not fully linear. The remaining non-linearity is concentrated at the bottom of the income distribution: low-income neighborhoods exhibit a higher income-price elasticity, while high-income neighborhoods are much closer to the middle-income benchmark. This pattern is primarily driven by chain retailers, whereas independent stores display much weaker heterogeneity.

We further examine the role of local market conditions and store-level characteristics. Differences in local competition and product assortment account for part of the observed heterogeneity, but do not eliminate it. Additional local factors also appear to contribute to pricing differences across markets. More broadly, income should be interpreted as a proxy for a bundle of local characteristics, including demand conditions, retail environments, and neighborhood attributes. Our objective is not to identify a causal effect of income per se, but to document how prices vary systematically across markets with different income levels and to characterize the margins underlying these differences. This approach relates to the decomposition of composition effects and price differences emphasized in previous work. Our setting focuses on within-city variation, allowing us to examine how these margins interact with retail structure and neighborhood characteristics.

Our paper contributes to this literature in three ways. First, we study the relationship between income and retail prices within a single city using highly disaggregated store–product data. This allows us to compare prices at different levels of aggregation—across all goods, within categories, within identical products, and within store–product pairs—and to separate composition effects from differences in the prices of comparable goods.

Second, we document that the aggregate income–price relationship largely reflects differences in product composition and retail structure. The estimated elasticity declines sharply as the comparison becomes more restrictive and becomes economically small for identical goods sold by the same retailer. This provides a unified interpretation of seemingly conflicting findings in the literature, showing that different results arise from comparing different objects.

Third, we show that the remaining heterogeneity is not uniform across the income distribution. The non-linearity is concentrated at the bottom of the distribution and is driven primarily by chain retailers. This highlights the role of retail organization and local market structure as key margins behind spatial price differences across neighborhoods.

Rather than providing a structural model of retail pricing, the paper aims to document a set of empirical regularities that clarify the mechanisms underlying spatial price differences in urban markets. While these mechanisms account for part of the observed heterogeneity, they do not fully explain it, suggesting that additional local factors also contribute to pricing differences across markets.

The remainder of the paper is organized as follows. Section 2 reviews the literature. Section 3 describes the data. Section 4 presents the empirical strategy and main results. Section 5 examines heterogeneity and related mechanisms. Section 6 concludes.

2 Related Literature

This paper relates to three strands of literature. First, it contributes to the literature on retail price setting across local markets. Standard theory highlights both the potential gains and costs of price discrimination across locations (Dobson and Waterson, 2005). Empirically, however, several studies show that large retailers, especially chains, often use pricing policies that are closer to uniform pricing than to fully flexible local price discrimination. This pattern has been documented for the United States (Nakamura and Steinsson, 2008; DellaVigna and Gentzkow, 2019; Adams and Williams, 2019; Hitsch, Hortaçsu, and Lin, 2021), for Argentina (Daruich and Kozlowski, 2023), and for Uruguay (Borraz and Zipitría, 2024). These pricing policies matter for interpreting spatial price differences: uniform pricing can compress price variation within chains, but it can also have distributional consequences when demand conditions, competitive pressure, or access to alternative stores differ across neighborhoods.

A related body of work emphasizes that retail markets are spatially segmented because consumers incur travel costs and do not freely arbitrage prices across locations. Eizenberg, Lach, and Oren-Yiftach (2021) show that residents of non-affluent peripheral neighborhoods in Jerusalem tend to shop in their home neighborhoods despite relatively high local prices, which they interpret as evidence of spatial frictions. Similar evidence for Uruguay points in the same direction: surveys show that most shopping trips are short, often starting from home, with many consumers walking only a few blocks to the store, although larger purchases involve longer trips (Comisión de Promoción y Defensa de la Competencia, 2022). This evidence motivates our focus on neighborhood-level prices: if retail markets are local, neighborhood income can be relevant for the demand environment faced by nearby stores.

Second, the paper addresses the long-standing literature on whether poorer consumers pay more. Since Caplovitz (1963), this question has generated mixed empirical findings. Some papers find that low-income households pay lower prices for the goods they purchase because they search more intensively, buy on sale, choose cheaper outlets, or adjust product choices (Aguiar and Hurst, 2007; Broda, Leibtag, and Weinstein, 2009; Beatty, 2010; Blow and Leicester, 2012). Other studies find that poorer consumers may face higher prices due to limited access to low-cost stores, smaller package sizes, liquidity constraints, or a weaker ability to take advantage of quantity discounts (Goodman, 1968; Chung and Myers, 1999; Rao, 2000; Attanasio and Frayne, 2006; Gibson and Kim, 2013).

A key issue in this literature is that different papers compare different objects. Some studies examine transaction prices paid by households, others study unit values, shopping baskets, or prices available at local stores. These distinctions matter because the prices households pay reflect store choice, timing, promotions, package size, and product selection. By contrast, our paper studies posted and transaction-based store prices across neighborhoods. We therefore do not estimate whether households with different incomes pay different prices for their realized purchases. Instead, we study whether the prices available in local retail environments vary systematically with neighborhood income, and whether this relationship reflects price differences for identical goods or differences in product and store composition.

Third, the paper relates to the literature on spatial cost-of-living differences, product availability, and local demand. A central insight in this literature is that observed price differences across locations need not reflect different prices for identical goods. They may instead reflect variation in product availability, store formats, consumption baskets, and quality. Handbury and Weinstein (2015) and Handbury (2021) show that product availability and non-homothetic demand are central for measuring spatial price differences. Relatedly, Jaravel (2021) emphasizes that product innovation and quality upgrading may differentially affect households across the income distribution. Local demand can also affect markups: Stroebel and Vavra (2019) shows that increases in local house prices raise retail prices, consistent with a demand-driven markup channel.

This perspective is directly related to our empirical strategy. We compare prices at several levels of aggregation: across all products, within categories, within identical products, and within store-product pairs. The sharp decline in the estimated income-price elasticity as the comparison becomes more restrictive suggests that much of the raw relationship between income and prices operates through product composition, store composition, and local retail structure, rather than through large price differences for identical goods sold by the same retailer.

Our paper is also related to recent work on neighborhood change and retail markets. Borraz, Carozzi, González-Pampillón, and Zipitría (2024) show that residential development in Montevideo affected both retail prices and product variety, highlighting that local income and neighborhood change can shape retail outcomes through multiple margins, including entry, markups, and assortment. This is close in spirit to our interpretation: neighborhood income is not only a demand shifter, but also a proxy for a broader bundle of local characteristics that affect retail structure and product availability.

Relative to this literature, our contribution is threefold. First, we study neighborhood income-price elasticities within a single city using highly disaggregated store-product data. Second, we cover both chain and independent retailers, which allows us to examine how the income-price relationship differs across retail organizations. Third, we document an asymmetric form of retail segmentation: the main non-linearity is concentrated at the bottom of the neighborhood income distribution, rather than reflecting a smooth premium for high-income neighborhoods across the entire distribution. The evidence therefore points to retail structure, local competition, and product availability as central margins behind spatial price differences, while price differences for identical goods appear limited.

3 Data

To study how retail prices vary with neighborhood income, we combine two complementary grocery price datasets with detailed neighborhood-level income information.

Our primary source is a dataset collected by the General Directorate of Commerce of the Uruguayan Ministry of Economy and Finance (*Dirección General de Comercio*, hereafter DGC). This dataset contains daily posted prices for up to 154 supermarket products across all major retail chains and a subset of independent stores, spanning April 2007 to December 2022. Products are selected to be comparable across stores and correspond primarily to the leading brands within each category.[1] After discarding non-homogeneous products, the final sample includes 128 comparable products.[2]

We complement this information with scanner data collected by Scanntech, a private firm that gathers transaction-level data for medium- and small-sized grocery stores. Prices in this dataset are constructed as unit values, defined as total sales divided by quantities, at the product-store-month level. We requested the same products as in the DGC database to ensure comparability across sources. The Scanntech data cover January 2013 to May 2022, with a short gap between January and March 2016.

The two datasets provide a useful trade-off. The DGC data offer dense product-level and time-series variation for a smaller set of larger retailers, while Scanntech expands coverage to a broader cross-section of smaller stores. The details of the database construction and manipulation are provided in Appendix B.

Our period of analysis is January 2013 through May 2022, excluding months with missing Scanntech data. To ensure comparability across sources, we aggregate DGC daily prices to the monthly level and deflate each product-store price using the monthly Consumer Price Index (CPI). Both price datasets include store location and the number of cashiers, which we use as a proxy for store size.

To capture local purchasing power, we merge the price data with neighborhood-level income constructed from household survey data provided by the *Instituto Nacional de Estadística* (INE). Since residents tend to shop close to home, neighborhood-level prices are a natural proxy for the prices faced by local consumers (Eizenberg, Lach, and Oren-Yiftach, 2021). Although the price data cover the whole country, reliable monthly income

[1] This database has been used previously in Borraz and Zipitría (2012) to characterize price rigidity, in Borraz, Cavallo, Rigobon, and Zipitría (2016) and Borraz and Zipitría (2022) to study price convergence, and in Borraz, Carozzi, González-Pampillón, and Zipitría (2024) to analyze demand shocks and stores' price responses.

[2] Appendix B details the product selection procedure.

measures at the neighborhood level can only be constructed for Montevideo. We therefore restrict the analysis to Uruguay's capital city, which allows us to use a consistent definition of local markets and to exploit meaningful cross-neighborhood variation in income.

Montevideo has 62 neighborhoods. We construct reliable monthly income measures for 49 of them.[3] The excluded neighborhoods represent 16.5% of the city's population and account for 12% of stores and 11% of price observations. We restrict the price databases to the 49 neighborhoods included in the income database.

Income is measured net of taxes and constructed as a six-month rolling average of household income within each neighborhood. This smoothing reduces sampling variability due to limited observations in some neighborhood-month cells and mitigates the effect of outliers. We further address extreme values and missing observations by applying an outlier detection procedure and interpolating income when necessary. Details of this procedure are provided in Appendix C. Finally, we deflate neighborhood average income using the monthly CPI.

Importantly, despite differences in how prices are measured—posted prices in DGC and transaction prices in Scanntech—both sources yield nearly identical price levels when matched at the store-product-month level. Some stores appear in both databases, allowing us to manually match stores using their addresses and compare price measures directly. We identify 98 stores in the DGC database and 108 in Scanntech as duplicates. The difference reflects changes in ownership for some Scanntech stores, which were entered as new stores.

For the overlapping stores, the distribution of price differences is highly concentrated around zero. The median log price difference is zero, and 37% of matched observations have identical prices in both sources. The interquartile range is approximately 0.18%, indicating very small differences between posted and transaction prices. Regressions of log-deflated prices on a Scanntech dummy, controlling for time, store-pair, product, and store-pair-by-product fixed effects, also show no statistically significant price differences

[3]The list of neighborhoods included and excluded is reported in Appendix C.

across sources. These results support the use of both datasets in a common price analysis. Descriptive statistics for the full databases and for duplicated stores are reported in Appendix A.

Because some stores appear in both sources, we define our baseline database by retaining only the DGC observation for duplicated stores. To characterize nonlinearities in the relationship between prices and income, we group neighborhoods into four quartiles based on the monthly income distribution. Our main analysis focuses on comparing the first and fourth quartiles, which provides a transparent contrast between low- and high-income markets and maximizes statistical power.

Figure 1: Store Location in Montevideo.

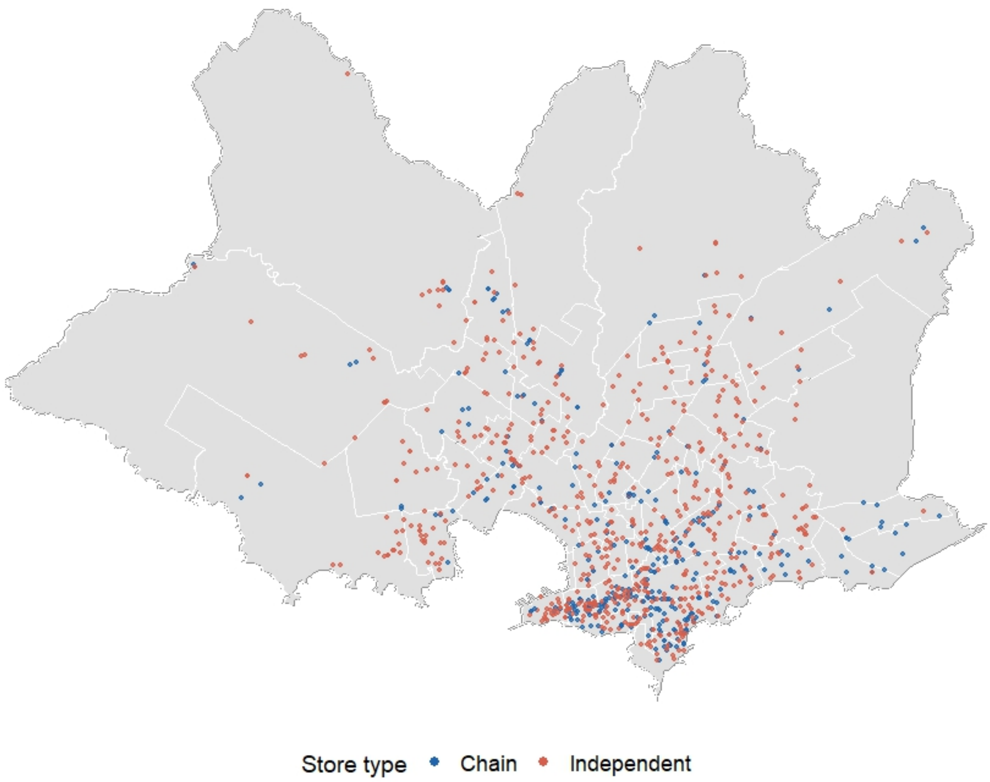


Figure 1 shows the spatial distribution of Montevideo stores, categorized as chain or independent. Most stores are located in the south-central part of the city, the most densely populated area.

Table 1 shows sharp differences in market conditions across income quartiles. Average

Table 1: Summary Statistics

Sample Period	01/2013	05/2022
Number of Markets (location)	49	
Number of Products	128	
Number of Categories	44	
Number of Stores	1,073	
Number of Chains	21	
Number of Observations	3,253,651	

	Quartile 1		Quartile 4	
	Mean	St. D.	Mean	St. D.
CPI Adjusted Income	3,938	954	9,841	2,746
CPI Adjusted Log Price	3.2305	0.5617	3.3109	0.5427
Number of Cashiers	2.31	2.82	2.47	3.21
Number of Products	44.36	27.49	48.36	26.36
Number of Stores	108.47	21.66	197.12	35.20
Number of Observations	615,940		1,224,390	

Source: Authors' calculations based on DGC and Scanntech data.
Notes: The upper panel reports overall sample coverage. The lower panel compares summary statistics for neighborhoods in the first and fourth income quartiles. The database covers January 2013 to May 2022, excluding January to March 2016, when Scanntech information is unavailable. The sample combines DGC and Scanntech data; when a store appears in both sources, the DGC observation is retained in the baseline database. Real income is measured at the neighborhood-month level; log deflated prices at the observation level; cashiers at the store level; products as the average number of products per store, first averaging within store over time and then across stores within each quartile; and stores as the average number of stores per month within each quartile. The number of observations is the total number of price observations in each quartile. CPI base year is 2022. Prices are expressed in April 2007 pesos, and income is expressed in December 2010 pesos.

real income in Quartile 4 markets is about 2.5 times that in Quartile 1 markets. Average log prices are also higher in richer neighborhoods, but the price gap is small relative to the underlying income difference. Price dispersion, measured by the standard deviation of log prices, is very similar across the two quartiles. This large variation in income across neighborhoods, combined with relatively small differences in prices, provides the key source of variation exploited in our empirical analysis.

Market structure differs more visibly across neighborhoods. Quartile 1 markets have fewer stores, somewhat fewer products, and slightly smaller stores on average than Quartile 4 markets. They also account for fewer price observations overall. These patterns suggest a smaller market scale in poorer neighborhoods, although the table alone is not sufficient to characterize retail density or effective market access.

Taken together, these patterns suggest that while income differences across neighborhoods are large, price differences are comparatively small. This raises the question of whether observed price differences reflect variation in product and store composition or genuine differences in the prices of comparable goods. This question is central to both the literature on spatial price differences, which emphasizes the role of consumption composition and product availability (Handbury and Weinstein, 2015; Handbury, 2021), and a long-standing literature on whether low-income households face higher prices systematically, which has reached mixed conclusions depending on the object of comparison (Goodman, 1968; Attanasio and Frayne, 2006; Broda, Leibtag, and Weinstein, 2009; Beatty, 2010; Gibson and Kim, 2013). The next section addresses this question by estimating how prices vary with neighborhood income after progressively accounting for product, store, and market characteristics.

4 Empirical Analysis

The descriptive evidence in Section 3 shows that income differences across neighborhoods are large, while observed price differences are comparatively small. Section 2 highlights two possible explanations: prices for identical goods may differ across local markets, or observed differences may reflect product and retailer composition.

We use the detailed product- and store-level variation in our data to decompose the relationship between prices and neighborhood income. Specifically, we compare prices at different levels of aggregation—across all products, within categories, within identical products, and within store-product pairs—so that the coefficient on income has a distinct economic interpretation in each specification.

Each set of fixed effects changes the economic object identified by the income coefficient. With time fixed effects only, the estimated elasticity captures cross-neighborhood differences in average prices within a month, combining variation in product composition, store types, and product quality. Adding category fixed effects restricts the comparison to similar groups of goods in the same month and isolates within-category price differences across neighborhoods. Introducing product fixed effects further narrows the comparison to identical goods sold at the same point in time, thereby removing differences in product composition and quality and isolating spatial price variation for homogeneous products. Finally, including both product and store fixed effects shifts the identifying variation to within-store, within-product price changes over time, capturing whether prices respond to changes in local income within a given retail environment.

The main equation is:

$$\log p_{ist} = \beta_1 \log y_{mt} + \beta_2 (\log y_{mt})^2 + \alpha_{ics} + \delta_t + \gamma_{Scann} + \varepsilon_{ist} \tag{1}$$

where p_{ist} is the price of product i in store s at time t, y_{mt} is real income in market m at time t, α_{ics} denotes several different fixed effects—α_i for product fixed effects, α_c for category fixed effects, α_s for store fixed effects—, δ_t denotes time fixed effects, γ_{Scann} is a dummy variable that have a value one if the data is from Scanntech database, and ε_{ist} is an error term. The specification includes income in both linear and quadratic terms.

We estimate this equation under alternative sets of fixed effects corresponding to the comparison margins described above. A decline in the income coefficient across these specifications indicates that part of the overall income-price elasticity is associated with differences in product or retailer composition across markets, while the remaining coefficient captures price differences among more comparable goods and retail environments. We cluster standard errors at the neighborhood-market and category levels. Table 2 reports the results.

Table 2: Income Elasticity Estimation.

Dependent Variable:	Log Deflated Price							
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Log Real Income	0.109***	0.067***	0.044***	0.021	-0.096***	0.024***	0.101***	-0.357
	(0.001)	(0.010)	(0.005)	(0.027)	(2.67×10^{-5})	(0.001)	(0.0002)	(0.228)
Scanntech Database	0.013	0.029**	0.020**		0.013	0.029**	0.020**	
	(0.043)	(0.012)	(0.008)		(0.043)	(0.014)	(0.008)	
(Log Real Income)^2					0.012***	0.002***	-0.003***	0.021*
					(0.0002)	(0.0007)	(0.0003)	(0.012)
Fixed-effects								
Time	Yes	Yes	Yes		Yes	Yes	Yes	
Category		Yes				Yes		
Product			Yes	Yes			Yes	Yes
Store				Yes				Yes
Fit statistics								
Observations	3,253,651	3,253,651	3,253,651	3,253,651	3,253,651	3,253,651	3,253,651	3,253,651
R^2	0.00601	0.88389	0.94674	0.94860	0.00603	0.88389	0.94674	0.94864
Within R^2	0.00418	0.01781	0.01634	0.00127	0.00420	0.01781	0.01637	0.00192

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. The dummy *Scanntech Database* equals one for observations from the Scanntech database and zero for observations from the DGC database. The baseline database combines DGC and Scanntech data; when a store appears in both sources, the DGC observation is retained. Columns (1)–(4) include income only in linear form, while columns (5)–(8) add a quadratic income term. The specifications progressively introduce fixed effects for time, category, product, and store, as indicated in the table. Standard errors are clustered at the category and market levels.

Table 2 reports a clear positive association between market income and prices across all linear specifications. In Column (1), which only absorbs common time effects, a one-standard-deviation increase in log real income (0.423) is associated with an estimated 4.6% increase in prices. This first specification captures broad cross-market differences in observed prices, combining variation in product composition, store types, and product quality. Once the comparison is restricted to products within the same broad category in the same month, the implied effect falls to 2.8% (Column 2). Restricting the comparison further to identical products across markets reduces the implied price difference to 1.8%

(Column 3), and adding store fixed effects lowers it to 0.2% (Column 4), so that the remaining variation comes from the same product sold by the same retailer over time.

This monotonic decline shows that the income-price elasticity is strongest in broad cross-market comparisons and becomes very small when the analysis is restricted to identical products sold by the same retailer. This implies that the positive association between income and prices largely reflects differences in consumption and retail environments across markets, rather than systematic price differences for identical goods. Instead, most of the income-price elasticity operates through differences in store composition across markets and in product differences within categories. This pattern is consistent with the idea that spatial price differences largely reflect composition effects, as emphasized in the literature.

This interpretation is also consistent with the observed dispersion in prices. There is more price variation within categories than within products: the average standard deviation of log real prices is 0.181 within categories and 0.135 within products. This reinforces the idea that the relevant margin of adjustment is substitution across products within categories, rather than price variation for identical goods. In other words, the price data point to product composition within categories as a more important source of cross-market price differences than variation in prices for identical goods. The additional decline from Column (3) to Column (4) indicates that part of the aggregate income-price relationship also reflects differences in retailer composition across markets, not only differences across goods.

Taken together, these results clarify the sources of the aggregate income-price relationship. A substantial share of the raw association between income and prices reflects differences in product composition and retail structure across markets, rather than large price differences for identical goods sold in the same outlet. As a result, the scope for price differences across neighborhoods lies primarily in store choice and in substitution across products within categories, rather than in differential pricing of tightly defined goods. This implies that the scope for systematic price differences for identical goods across neighborhoods is very limited.

Columns (5) to (8) show that this relationship is not always well described by a linear specification. In the quadratic models, the income-price elasticity varies with income according to $\partial \log p/\partial \log y = \beta_1 + 2\beta_2 \log y$. In Columns (5) and (6), the squared term is positive, implying a convex relationship in which the elasticity rises with income over the observed range. In Columns (7) and (8), by contrast, the squared term is negative, implying a concave relationship once the comparison is restricted to identical products, and especially to identical products within the same store. In these specifications, prices still tend to increase with income, but at a decreasing rate; in Column (8), the elasticity becomes very small at high income levels.

Overall, the results show that the income-price elasticity is positive but highly sensitive to the comparison: it is strongest for broad cross-market price differences and weakest when the analysis focuses on identical goods sold by the same retailer. The fact that the curvature changes as the comparison is restricted suggests that nonlinearities in the income-price relationship are primarily driven by composition effects at broader levels of aggregation, while price differences for identical goods remain limited across the income distribution.

Table 14 in Appendix A replicates Table 2 using an alternative version of the database in which duplicated stores are assigned to Scanntech rather than DGC. The results are very similar to the baseline estimates. In the linear specifications, the coefficient on log real income remains positive and declines as the comparison is restricted from broad market prices to within-category, within-product, and within-store variation. The nonlinear specifications deliver the same general pattern. Thus, the alternative database does not change the main conclusion: the positive association between income and prices is robust, but much of it reflects composition across categories, products, and stores rather than large price differences for identical goods sold by the same retailer.

Having established that these patterns are robust to the treatment of duplicated stores, we now turn to the form of the non-linearity in the income-price relationship. We use the quadratic specifications in Table 2 as a preliminary characterization of this non-linearity, but focus on the quartile specification for interpretation. To examine the pattern more transparently, we group markets into income quartiles and interact log real income with indicators for the bottom and top quartiles, using middle-income markets (Quartiles 2 and 3) as the omitted category.

We now estimate the following specification:

$$\begin{aligned} \log p_{ist} = \beta_1 \log y_{mt} + \theta_1 Q1_{mt} + \theta_4 Q4_{mt} + \phi_1(\log y_{mt} \times Q1_{mt}) \\ + \phi_4(\log y_{mt} \times Q4_{mt}) + \alpha_{ics} + \delta_t + \gamma_{Scann} + \varepsilon_{ist}. \end{aligned} \tag{2}$$

where $Q1_{mt}$ and $Q4_{mt}$ are indicators for markets in the bottom and top income quartiles, respectively, and Quartiles 2 and 3 are omitted. The coefficient β_1 measures the income-price elasticity for middle-income markets, while ϕ_1 and ϕ_4 capture how that elasticity differs in Quartiles 1 and 4. Likewise, θ_1 and θ_4 measure level differences in prices relative to middle-income markets, conditional on the included controls. As in the previous table, we vary the set of fixed effects so that the coefficients can be interpreted under different comparison margins: across all observed prices within a month, within broad product categories, across identical products, and within identical products sold by the same retailer.

This specification complements the quadratic regressions in Table 2 by allowing the income-price relationship to vary across segments of the income distribution. Rather than imposing a single polynomial form, the quartile-interaction model more directly shows where non-linearity arises and whether it persists in narrower product- and retailer-level comparisons. Table 3 reports the estimates.

Table 3 shows that the non-linearity in the income-price relationship is largely concentrated at the bottom of the distribution. This pattern is consistent with the nonlinearities suggested by the quadratic specifications, but provides a clearer interpretation of where they arise. Across all specifications, Quartile 1 exhibits a higher income-price elasticity

Table 3: Income Quartile Interaction Estimation.

Dependent Variable:	Log Deflated Price		
Model:	(1)	(2)	(3)
Variables			
log(real_income)	0.128***	0.074***	0.049***
	(0.006)	(0.012)	(0.007)
Quartile 1	-0.255***	-0.184***	-0.162***
	(0.002)	(0.022)	(0.0005)
Quartile 4	-0.193***	-0.079***	-0.014***
	(0.0009)	(0.003)	(0.0005)
sourcescanntech	0.013	0.029**	0.020**
	(0.044)	(0.012)	(0.008)
Log Real Income x Q. 1	0.033***	0.023***	0.020***
	(0.0005)	(0.002)	(0.0004)
Log Real Income x Q. 4	0.020***	0.008***	0.001***
	(0.0003)	(0.0006)	(0.0002)
Fixed-effects			
Time	Yes	Yes	Yes
Category		Yes	
Product			Yes
Fit statistics			
Observations	3,253,651	3,253,651	3,253,651
R^2	0.00614	0.88392	0.94675
Within R^2	0.00431	0.01801	0.01656

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. Quartile 1 and Quartile 4 are dummy variables for neighborhoods in the lowest and highest income quartiles, respectively. The omitted group corresponds to the middle-income quartiles. The variables *Log Real Income* $\times$ *Q*.1 and *Log Real Income* $\times$ *Q*.4 interact log real income with the first and fourth quartiles. The dummy *Scanntech Database* equals one for observations from the Scanntech database and zero for observations from the DGC database. The baseline database combines DGC and Scanntech data; when a store appears in both sources, the DGC observation is retained. Time, category, and product fixed effects are included as indicated in the table. Standard errors are clustered at the category and market levels.

than the omitted middle-income group, whereas Quartile 4 remains much closer to that benchmark. Thus, the quartile specification confirms the earlier evidence while providing a clearer interpretation: the non-linearity is better characterized as a bottom-quartile premium in the income-price elasticity than as a smooth symmetric curvature.

In Model (1), which only includes time fixed effects, the estimated income-price elasticity is 0.128 for middle-income markets, 0.162 for Quartile 1, and 0.148 for Quartile 4. Interpreted in elasticity terms, a 10% increase in real income is associated with an increase in prices of about 1.3% in middle-income markets, 1.6% in Quartile 1, and 1.5% in Quartile 4. Evaluated at the mean log income of each group, the implied price difference relative to middle-income markets is about +2.2% for Quartile 1 and −0.4% for Quartile 4. Since this specification does not control for differences in categories, products, or retailers, it captures broad cross-market differences in observed prices.

Once category fixed effects are introduced in Model (2), the comparison is restricted to goods within the same broad category. Under this margin, the level differences become much smaller: approximately +0.3% for Quartile 1 and −0.3% for Quartile 4. The corresponding elasticities also decline, to 0.073 for middle-income markets, 0.096 for Quartile 1, and 0.081 for Quartile 4. In elasticity terms, a 10% increase in income is associated with price increases of about 0.7%, 1.0%, and 0.8%, respectively. Thus, part of the raw non-linearity reflects differences in product composition across broad classes of goods, but the bottom quartile still shows the strongest price response to income.

The same pattern continues in Model (3), which includes product fixed effects and therefore compares identical products across markets. At this level, the estimated price differences relative to middle-income markets are nearly zero: Quartile 1 is only about +0.1% above the omitted group, while Quartile 4 is about −0.3% below it. The corresponding elasticities are 0.049 for middle-income markets, 0.069 for Quartile 1, and 0.050 for Quartile 4. This implies that a 10% increase in income is associated with price changes of only 0.5%, 0.7%, and 0.5%, respectively. Once attention is restricted to the same good, most of the remaining non-linearity appears in the slope rather than in persistent differences in price levels.

Taken together, these results reinforce the message of Table 2 and clarify its economic content. As the comparison moves from all observed prices to goods within the same category, to identical products, the estimated level differences across income groups shrink sharply and become negligible. This indicates that price differences across income groups arise primarily from variation in store environments and product composition, rather than from differences in the prices of identical goods.

At the same time, the remaining non-linearity is concentrated in the slope: low-income markets remain more income-elastic than middle- or high-income markets. This is economically relevant because Quartile 1 combines higher elasticity with, in the less-restrictive specifications, higher prices relative to the middle-income benchmark. In other words, low-income markets tend to exhibit both slightly higher price levels in broader comparisons and greater price sensitivity to income within those markets. Relative to the quadratic specifications, the quartile-interaction results therefore show more clearly that the aggregate non-linearity is driven mainly by the lower tail of the income distribution.

5 Heterogeneity in the Income-Price Elasticity

The results in the previous section show that the income-price elasticity becomes much weaker once the comparison is restricted to identical goods, and that the remaining non-linearity is concentrated in low-income markets. A natural next step is to examine whether this heterogeneity is related to differences in the types of retailers operating across markets. We first examine whether the heterogeneity in the income-price relationship is driven by differences in retail organization. We then assess whether local market conditions and store-level characteristics—such as competition, size, and assortment—account for part of the remaining variation.

Store ownership provides a central margin of heterogeneity in the income-price elasticity. Chain retailers may rely on centralized pricing, standardized product policies, and broader organizational constraints, whereas independent stores may adjust prices more flexibly in response to local market conditions. This distinction is consistent with the evidence in DellaVigna and Gentzkow (2019), who show that large U.S. retail chains charge nearly uniform prices across stores despite substantial variation in local demographics and competition, reflecting limited price adjustment to local conditions.

Motivated by this evidence, Table 4 splits the sample between chain and independent stores and re-estimates the quartile-interaction regressions separately for each group. This allows us to assess whether the non-linearity documented above is a general feature of retail pricing or is instead concentrated in a specific segment of the sector.

Table 4: Income Quartile Interaction Estimation. Sample: Chains

Dependent Variable:	Log Deflated Price			
Chain Type	Chains	Indep.	Chains	Indep.
Model:	(1)	(2)	(3)	(4)
Variables				
Log Real Income	0.045***	0.072***	0.031***	0.049***
	(0.011)	(0.015)	(0.009)	(0.011)
Quartile 1	-0.306***	-0.100	-0.288***	-0.069
	(0.092)	(0.099)	(0.073)	(0.075)
Quartile 4	-0.026	-0.151*	0.028	-0.067*
	(0.064)	(0.076)	(0.040)	(0.039)
Log Real Income x Q. 1	0.036***	0.013	0.034***	0.009
	(0.010)	(0.012)	(0.008)	(0.009)
Log Real Income x Q. 4	0.002	0.017*	-0.004	0.007
	(0.007)	(0.009)	(0.004)	(0.005)
Fixed-effects				
Category	Yes	Yes		
Time	Yes	Yes	Yes	Yes
Product			Yes	Yes
Fit statistics				
Observations	1,520,240	1,779,016	1,520,240	1,779,016
R^2	0.87571	0.89845	0.94547	0.95318
Within R^2	0.00478	0.01516	0.00481	0.01376

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. Quartile 1 and Quartile 4 are dummy variables for neighborhoods in the lowest and highest income quartiles, respectively; the omitted group corresponds to the middle-income quartiles. The variables *Log Real Income* $\times$ *Q.* 1 and *Log Real Income* $\times$ *Q.* 4 interact log real income with the first and fourth quartiles. The sample combines DGC and Scanntech data; when a store appears in both sources, the DGC observation is retained in the baseline database. Columns (1) and (3) are estimated on chain stores only, whereas columns (2) and (4) are estimated on independent stores only. Columns (1) and (2) include category and time fixed effects, while columns (3) and (4) include product and time fixed effects. Standard errors are clustered at the category and market levels.

Table 4 shows that the aggregate quartile pattern masks substantial heterogeneity across store types. In the full-sample specification with time and category fixed effects, the estimated income-price elasticities were 0.073 for middle-income markets, 0.096 for Quartile 1, and 0.081 for Quartile 4. Once the sample is split, chain stores display a much sharper non-linearity. Their implied elasticity is 0.045 for middle-income markets, 0.081 for Quartile 1, and 0.047 for Quartile 4. Thus, among chains, the bottom quartile is markedly more income-elastic than both the omitted middle-income group and the top quartile.

Independent stores, by contrast, exhibit a much flatter pattern: the corresponding elasticities are 0.072, 0.085, and 0.089. Their baseline elasticity is very close to the full-sample estimate, and both Quartile 1 and Quartile 4 lie only modestly above it. This indicates that the strong bottom-quartile premium observed in the aggregate data is not a general feature of all retailers, but is much more pronounced among chains.

The same conclusion becomes even clearer once the comparison is restricted to identical products. In the full-sample specification with time and product fixed effects, the estimated income-price elasticities were 0.049 for middle-income markets, 0.069 for Quartile 1, and 0.050 for Quartile 4. For chain stores, the corresponding elasticities are 0.031, 0.065, and 0.027. Hence, within chains, the bottom quartile remains much more income-elastic than both the omitted group and the top quartile, and Quartile 4 is even slightly less elastic than middle-income markets.

Among independent stores, instead, the elasticities are 0.049, 0.058, and 0.056, which are much closer to one another. Once prices are compared for the same product, independents display only limited heterogeneity across the income distribution, whereas the chain segment continues to show a pronounced bottom-quartile pattern.

Table 15 in Appendix A shows that this pattern is robust to including chain fixed effects. This specification compares prices within the same retail chain, netting out persistent differences in price levels across chains. The bottom-quartile differential remains positive and statistically significant: it is 0.020 in the category fixed-effects specification and 0.018 in the product fixed-effects specification. The corresponding differentials for

Quartile 4 are much smaller, at 0.007 and 0.001. At the same time, the baseline income elasticity for middle-income markets becomes close to zero, suggesting that part of the chain-store pattern in Table 4 reflects differences in the composition of chains across neighborhoods.

Overall, these results indicate that the non-linear income-price elasticity documented in the full sample is especially pronounced among chain retailers. For chains, Quartile 1 remains distinctly more income-elastic, while Quartile 4 stays close to, or below, the middle-income benchmark. For independent stores, by contrast, the elasticity is much more homogeneous across quartiles, especially once the comparison is restricted to identical products. Robustness checks with chain fixed effects show that the bottom-quartile pattern is not driven solely by differences in chain composition across neighborhoods, although such differences account for part of the baseline chain-store estimates.

Economically, this suggests that the aggregate non-linearity is linked not only to neighborhood market conditions, but also to the organizational structure of retailing. The sharper bottom-quartile pattern appears to be concentrated in the chain segment, whereas independent stores show a much flatter relationship between prices and income.

5.1 Store Competition

The previous results indicate that the non-linear income-price elasticity is concentrated primarily among chain retailers. This points to the organizational structure of retailing as a central margin of heterogeneity. However, retail organization is not the only relevant dimension. Local market conditions may also shape how prices vary with neighborhood income. In particular, neighborhoods differ in the density of competing stores, in population density, and in persistent local characteristics that may affect retail costs, demand, and competitive pressure.

This margin is closely related to the literature on local retail pricing and zone pricing. Retail chains often do not set fully store-specific prices, even when local demand and competitive conditions differ across markets (DellaVigna and Gentzkow, 2019). At the same time, the literature on zone pricing emphasizes that local competition can affect the gains from finer geographic pricing and the distribution of prices across markets (Adams and Williams, 2019). In our context, this suggests that the stronger income-price elasticity observed in low-income markets may partly reflect differences in local competitive density rather than income alone.

We therefore examine whether the heterogeneity in the income-price elasticity is related to local competitive conditions. A natural possibility is that part of the stronger elasticity observed in low-income markets reflects differences in the density of competing stores across neighborhoods. To assess this channel, we augment the quartile-interaction specification by adding a measure of local store competition, defined as the number of competing stores each month per 1,000 inhabitants and square kilometer, and allow its association with prices to differ across income quartiles.

This specification is informative for two reasons. First, it shows whether prices are systematically lower in markets with greater local store density. Second, it indicates whether controlling for competition alters the quartile pattern documented above, in particular the stronger income-price elasticity observed in Quartile 1. If the quartile differences were largely driven by local competition, one would expect the estimated elasticity differentials across quartiles to shrink substantially once this variable is included. Table 5 reports the results.

Table 5 examines whether the heterogeneity in the income-price elasticity is related to differences in local store competition. Relative to the baseline quartile regressions, these specifications add a control for competitive density, measured as the number of competing stores per 1,000 residents per square kilometer, and interact this variable with the bottom and top income quartiles. This allows the price-competition relationship to differ across market-income groups.

In the specification with time and category fixed effects, the estimated income-price elasticity for middle-income markets falls from 0.073 in the baseline regression to 0.0616 once local competition is included. The corresponding elasticity for Quartile 1 declines

Table 5: Income Quartile Estimation with Local Competition Controls.

Dependent Variable:	Log Deflated Price	
Model:	(1)	(2)
Variables		
Log Real Income	0.0619***	0.0410***
	(0.0126)	(0.0069)
Local Store Competition	-0.0007***	-0.0005***
	(9.71×10^{-5})	(7.39×10^{-5})
Quartile 1	-0.1853***	-0.1666***
	(0.0276)	(4.93×10^{-5})
Quartile 4	-0.2178***	-0.1043***
	(0.0105)	(0.0003)
Log Real Income x Q. 1	0.0230***	0.0204***
	(0.0034)	(0.0004)
Log Real Income x Q. 4	0.0226***	0.0106***
	(0.0019)	(0.0003)
Local Store Competition × Quartile 1	0.0002	0.0002**
	(0.0003)	(9.27×10^{-5})
Local Store Competition × Quartile 4	0.0066	0.0041***
	(0.0055)	(0.0009)
Fixed-effects		
Category	Yes	
Time	Yes	Yes
Product		Yes
Fit statistics		
Observations	3,253,651	3,253,651
R^2	0.88331	0.94649
Within R^2	0.01285	0.01174

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. Quartile 1 and Quartile 4 are dummy variables for neighborhoods in the lowest and highest income quartiles, respectively; the omitted group corresponds to the middle-income quartiles. The variables *Log Real Income* × *Q.* 1 and *Log Real Income* × *Q.* 4 interact log real income with the first and fourth quartiles. *Local Store Competition* is defined as the number of competing stores in the neighborhood, normalized by population and area (stores per 1,000 residents per km^2). The variables *Local Store Competition* × *Quartile* 1 and *Local Store Competition* × *Quartile* 4 allow the effect of local competition to differ in the lowest- and highest-income quartiles. The sample combines DGC and Scanntech data; when a store appears in both sources, the DGC observation is retained in the baseline database. Column (1) includes category and time fixed effects, while column (2) includes product and time fixed effects. Standard errors are clustered at the category and market levels.

from 0.096 to 0.0848, while that for Quartile 4 changes only slightly, from 0.081 to 0.0841. Thus, controlling for local competition reduces the overall magnitude of the income-price elasticity and narrows the gap between middle-income and low-income markets. In this broader comparison, it also makes the elasticity in the top quartile closer to that of the bottom quartile.

A similar pattern emerges in the specification with time and product fixed effects, where prices are compared for identical goods across markets. In this case, the baseline elasticity for middle-income markets declines from 0.049 to 0.0407 after adding the competition control. The elasticity for Quartile 1 falls from 0.069 to 0.0613, whereas Quartile 4 remains essentially unchanged, moving from 0.050 to 0.0512. This indicates that local competitive density accounts for part of the stronger income-price elasticity previously observed in low-income markets, but it does not eliminate it.

The competition variable itself enters with a negative and significant coefficient in both specifications. For the omitted middle-income group, a higher density of competing stores is associated with lower prices: the coefficient is -0.0007 in the category fixed-effects regression and -0.0005 in the product fixed-effects regression. Economically, this means that higher local competition is associated with lower prices, though the effect is modest. Using the observed dispersion of the variable, a one-standard-deviation increase in competition in middle-income markets is associated with approximately 0.5% lower prices in the category specification and 0.4% lower prices in the product specification.

There is also some heterogeneity across quartiles in the competition effect. For Quartile 1, the interaction term is small in both columns and is statistically significant only in the product fixed-effects specification, suggesting that the negative relationship between competition and prices is, if anything, slightly weaker in low-income markets when identical products are compared. For Quartile 4, the interaction is positive in both columns and strongly significant in the product fixed-effects regression, implying that the price-reducing effect of competition is weaker in top-income markets. In the category specification, the Quartile 4 interaction is estimated imprecisely, so the result should be interpreted with caution.

These patterns are also consistent with the descriptive distribution of competitive density across markets. Competition is substantially stronger in low-income neighborhoods: the mean number of competing stores per 1,000 residents is 8.58 in Quartile 1, compared with 2.99 in middle-income markets and only 1.57 in Quartile 4. This suggests that part of the quartile heterogeneity documented in the baseline regressions reflects the fact that poorer markets are also more locally competitive.

Overall, controlling for local competition accounts for part of the previously estimated non-linearity, primarily by reducing the excess elasticity of the bottom quartile relative to the omitted middle-income group. However, it does not overturn the central result. Even after accounting for competitive density, low-income markets continue to display the highest income-price elasticity, especially once the comparison is restricted to identical products. Local competition explains part of the income-price relationship, but a substantial share of the heterogeneity remains even after controlling for this margin.

The aggregate income-price relationship, therefore, appears to reflect both observable differences in local market structure and additional local factors not fully captured by competition alone. At the same time, these results complement the evidence above that the strongest non-linearity is concentrated among chain retailers. Local competition and neighborhood characteristics help account for part of the pattern, but they do not fully explain the heterogeneity in the income-price elasticity, which remains closely linked to the organizational structure of retail markets.

5.2 Store-Level Factors

We next turn to store-level factors. The previous subsections show that heterogeneity in income-price elasticity is concentrated among chain retailers and is only partially explained by local competitive conditions and neighborhood-level factors. This raises the question of whether additional heterogeneity arises from differences in store characteristics, such as size and product assortment, that vary across markets.

This margin is related to a broader literature on retail structure and pricing, which emphasizes that differences in store format, scale, and assortment can affect both cost conditions and the set of goods available within stores (Basker, 2007; Handbury and Weinstein, 2015; Jaravel, 2021). Larger stores may benefit from economies of scale and lower procurement costs, while broader assortments may reflect differences in product availability and substitution opportunities across markets. In this context, we examine whether these store-level characteristics help account for the remaining heterogeneity in the income-price relationship.

Store size is a natural candidate mechanism behind the heterogeneity in the income-price elasticity. Larger stores may operate with lower unit costs and economies of scale, but may also differ in pricing strategies and market power. These differences may vary systematically across neighborhoods. To assess this margin, Table 6 augments the baseline quartile-interaction specification by adding *Log Number of Cashiers* as a proxy for store size, together with interactions between store size and the bottom and top income quartiles. This allows the association between store size and prices to vary across market-income groups, while preserving the same comparison structure as in the baseline quartile regressions.

The table is informative for two related questions. First, it shows whether controlling for store size reduces the estimated income-price elasticity across quartiles, which would suggest that part of the baseline pattern reflects differences in retailer scale. Second, it shows whether the relationship between store size and prices itself differs across the income distribution. In this way, the specification helps determine whether the stronger elasticity previously found in low-income markets is partly explained by differences in store size, or whether that result remains after conditioning on this margin of store heterogeneity.

Controlling for store size leaves the main quartile pattern largely unchanged. Relative to the baseline quartile regressions, Table 6 adds *Log Number of Cashiers* as a proxy for store size and allows its association with prices to vary across income quartiles. In the specification with time and category fixed effects, the estimated income-price elasticity for the omitted middle-income group falls from 0.073 in the baseline regression to 0.0674. The implied elasticity for Quartile 1 becomes $0.0674 + 0.0235 = 0.0909$, compared with 0.096 in the baseline, while the elasticity for Quartile 4 is $0.0674 + 0.0132 = 0.0806$,

Table 6: Income Quartile Estimation (Controlled by Store Size).

Dependent Variable:	Log Deflated Price	
Model:	(1)	(2)
Variables		
Log Real Income	0.0674***	0.0447***
	(0.0116)	(0.0068)
Log Number of Cashiers	-0.0067	-0.0053
	(0.0062)	(0.0044)
Quartile 1	-0.1946***	-0.1685***
	(0.0244)	(0.0011)
Quartile 4	-0.1068***	-0.0315***
	(0.0020)	(0.0004)
Log Real Income x Q. 1	0.0235***	0.0202***
	(0.0017)	(0.0009)
Log Real Income x Q. 4	0.0132***	0.0043***
	(0.0008)	(0.0004)
Log Number of Cashiers × Quartile 1	0.0051	0.0039
	(0.0094)	(0.0071)
Log Number of Cashiers × Quartile 4	-0.0159***	-0.0095***
	(0.0035)	(0.0019)
Fixed-effects		
Category	Yes	
Time	Yes	Yes
Product		Yes
Fit statistics		
Observations	3,299,256	3,299,256
R^2	0.88389	0.94681
Within R^2	0.01761	0.01614

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. Quartile 1 and Quartile 4 are dummy variables for neighborhoods in the lowest and highest income quartiles, respectively; the omitted group corresponds to the middle-income quartiles. The variables *Log Real Income*×*Q.* 1 and *Log Real Income*×*Q.* 4 interact log real income with the first and fourth quartiles. *Log Number of Cashiers* is the logarithm of the number of cashiers in the store and is used as a proxy for store size. The variables *Log Number of Cashiers* × *Quartile* 1 and *Log Number of Cashiers* × *Quartile* 4 allow the association between store size and prices to differ in the lowest- and highest-income quartiles. The sample combines DGC and Scanntech data; when a store appears in both sources, the DGC observation is retained in the baseline database. Column (1) includes category and time fixed effects, while column (2) includes product and time fixed effects. Standard errors are clustered at the category and market levels.

essentially unchanged relative to the previous estimate of 0.081. Thus, conditioning on store size modestly attenuates the stronger price-income elasticity at the bottom of the distribution, but leaves the main non-linearity intact.

The same conclusion emerges once the comparison is restricted to identical products. In the specification with time and product fixed effects, the elasticity for middle-income markets declines from 0.049 in the baseline regression to 0.0447. The corresponding elasticity for Quartile 1 is $0.0447 + 0.0202 = 0.0649$, down from 0.069, while the elasticity for Quartile 4 is $0.0447 + 0.0043 = 0.0490$, virtually identical to the earlier estimate of 0.050. Hence, even when identical goods are compared across markets, controlling for store size only modestly changes the magnitudes and does not alter the qualitative ranking across quartiles: Quartile 1 remains the most income-elastic group.

The direct association between store size and prices is weak on average. The coefficient on *Log Number of Cashiers* is small and statistically insignificant in both specifications, at -0.0067 in the category fixed-effects regression and -0.0053 in the product fixed-effects regression. For the omitted middle-income group, this implies that there is no economically meaningful average relationship between store size and prices. Since the regressor is in logs, a 10% increase in the number of cashiers is associated with only about 0.067% lower prices in the first specification and 0.053% lower prices in the second, both quantitatively negligible effects.

There is also limited descriptive variation in store size across quartiles. The mean number of cashiers is 2.48 in middle-income markets, 2.31 in Quartile 1, and 2.47 in Quartile 4, with standard deviations of 3.37, 2.82, and 3.21, respectively. These figures suggest that differences in average store size across income groups are modest, consistent with this variable's limited role in explaining the baseline quartile differences.

The interaction terms indicate that the association between store size and prices is not meaningfully different in low-income markets, but is more negative in high-income ones. The interaction *Log Number of Cashiers* $\times$ *Q. 1* is positive but statistically insignificant in both specifications, implying no clear difference relative to the omitted middle-income group. By contrast, the interaction with Quartile 4 is negative and significant in both columns, at -0.0159 and -0.0095. This means that in top-income markets, the total coefficient on store size becomes -0.0226 in the category comparison and -0.0148 in the product comparison. Interpreted proportionally, a 10% increase in the number of cashiers is associated with about 0.23% lower prices in the first case and 0.15% lower prices in the second. So the negative price-size relationship is economically more relevant in Quartile 4 than in other quartiles, though still modest in magnitude.

Overall, the table shows that store size explains only a limited part of the previously documented non-linearity in the income-price relationship. Controlling for cashier counts slightly reduces the estimated elasticity for middle-income markets and for Quartile 1, but does not overturn the central result: the bottom quartile continues to exhibit the highest income-price elasticity in both specifications, while Quartile 4 remains much closer to the middle-income benchmark. In short, differences in store size do not appear to be the main mechanism behind the stronger income-price elasticity observed in low-income markets.

We next examine whether the quartile pattern is related to differences in product assortment across stores. To do so, we augment the baseline quartile-interaction specification with *Sh. Store Products*, defined as the share of total products carried by the store, and allow its association with prices to vary across income quartiles. This variable captures a store-level dimension of assortment breadth rather than local market competition: a higher value indicates that the store offers a broader fraction of the overall product universe. This margin is closely related to the literature on product availability and non-homothetic demand, where differences in assortment across markets can generate differences in observed prices even when the prices of identical goods are similar (Handbury and Weinstein, 2015; Handbury, 2021; Jaravel, 2021). If part of the higher income-price elasticity in low-income markets is driven by narrower assortments, then controlling for *Sh. Store Products* should attenuate the quartile differences in estimated elasticities. Table 7 reports the results.

Table 7 examines whether the quartile pattern documented above is related to differences in product assortment across stores. Relative to the baseline quartile regressions, these specifications add a control for the store's product assortment share, *Sh. Store Products*,

Table 7: Income Quartile Estimation (Controlled by Store Assortment)

Dependent Variable:	Log Deflated Price	
Model:	(1)	(2)
Variables		
Log Real Income	0.062***	0.041***
	(0.011)	(0.006)
Sh. Store Products	-0.065**	-0.039**
	(0.028)	(0.019)
Quartile 1	-0.154***	-0.148***
	(0.012)	(0.003)
Quartile 4	-0.125***	-0.048***
	(0.006)	(0.0010)
Log Real Income x Q. 1	0.018***	0.017***
	(0.002)	(0.001)
Log Real Income x Q. 4	0.018***	0.008***
	(0.002)	(0.001)
Sh. Store Products $\times$ Quartile 1	0.021	0.009
	(0.020)	(0.019)
Sh. Store Products $\times$ Quartile 4	-0.068***	-0.044***
	(0.023)	(0.014)
Fixed-effects		
Category	Yes	
Time	Yes	Yes
Product		Yes
Fit statistics		
Observations	3,299,256	3,299,256
R^2	0.88442	0.94697
Within R^2	0.02213	0.01907

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. Quartile 1 and Quartile 4 are dummy variables for neighborhoods in the lowest and highest income quartiles, respectively; the omitted group corresponds to the middle-income quartiles. The variables *Log Real Income* $\times$ *Q.* 1 and *Log Real Income* $\times$ *Q.* 4 interact log real income with the first and fourth quartiles. *Sh. Store Products* is the store share of products, defined as the number of products available in the store divided by the total number of products observed in the sample in that month. The variables *Sh. Store Products* $\times$ *Quartile* 1 and *Sh. Store Products* $\times$ *Quartile* 4 allow the association between store product availability and prices to differ in the lowest- and highest-income quartiles. The sample combines DGC and Scanntech data; when a store appears in both sources, the DGC observation is retained in the baseline database. Column (1) includes category and time fixed effects, while column (2) includes product and time fixed effects. Standard errors are clustered at the category and market levels.

as defined above, and allow its association with prices to vary across income quartiles.

In the earlier specification with time and category fixed effects, the estimated income-price elasticities were 0.073 for middle-income markets, 0.096 for Quartile 1, and 0.081 for Quartile 4. After controlling for *Sh. Store Products* in Column (1), the corresponding elasticities become 0.0624 for middle-income markets, 0.0799 for Quartile 1, and 0.0801 for Quartile 4. Thus, adding the assortment control reduces the baseline elasticity and compresses the bottom-quartile premium, while leaving the top quartile essentially unchanged. The same pattern appears in the product fixed-effects specification. In the earlier regression, the estimated elasticities were 0.049 for middle-income markets, 0.069 for Quartile 1, and 0.050 for Quartile 4. After controlling for *Sh. Store Products* in Column (2), they become 0.0413, 0.0587, and 0.0491, respectively. Hence, part of the stronger income-price elasticity previously estimated for the bottom quartile appears to be associated with differences in assortment breadth, but the basic ranking across quartiles remains unchanged.

The direct association between assortment breadth and prices is negative in both specifications. In the omitted middle-income group, the coefficient on *Sh. Store Products* is -0.0648 in Column (1) and -0.0393 in Column (2), implying that stores carrying a broader share of products tend to have lower prices on average. The interaction between *Sh. Store Products* and Quartile 1 is positive but statistically insignificant in both columns, which suggests that this relationship is not clearly different in low-income markets. By contrast, the interaction with Quartile 4 is negative and significant in both cases, indicating that the negative association between assortment breadth and prices is substantially stronger in top-income markets.

This pattern is economically meaningful. Assortment breadth differs across quartiles: mean *Sh. Store Products* is 0.402 in middle-income markets, 0.363 in Quartile 1, and 0.417 in Quartile 4, with standard deviations of 0.217, 0.225, and 0.211, respectively. In middle-income markets, a one-standard-deviation increase in *Sh. Store Products* is associated with roughly 1.4% lower prices in the category specification and 0.9% lower prices in the product specification. In Quartile 4, the total coefficient on *Sh. Store Products* becomes -0.1326 in Column (1) and -0.0829 in Column (2), so that a one-standard-deviation increase in assortment breadth is associated with approximately 2.8%

and 1.7% lower prices, respectively. Thus, broader assortments are associated with lower prices everywhere, but especially in high-income markets.

Overall, controlling for the share of products carried by the store accounts for part of the previously estimated non-linearity, primarily by reducing the excess elasticity in the bottom quartile. At the same time, stores with broader assortments tend to charge lower prices, and this negative association is particularly strong in top-income markets. Therefore, assortment breadth accounts for part of the income-price relationship, but it does not overturn the main result: Quartile 1 remains more income-elastic than the omitted middle-income group, especially once the comparison is restricted to identical products.

Taken together, the evidence in this subsection shows that store-level characteristics account for part, but not all, of the heterogeneity in the income-price elasticity. Differences in store size play only a limited role: controlling for the number of cashiers leaves both the magnitude and the ranking of the elasticities across income groups largely unchanged.

By contrast, product assortment contributes more meaningfully to the observed patterns. Stores with broader assortments tend to have lower prices, and controlling for assortment reduces part of the excess elasticity observed in the bottom quartile. This is consistent with the idea that differences in product availability across stores generate differences in observed prices across markets, even when the prices of identical goods are similar.

However, these margins do not overturn the main results. Even after accounting for store size and assortment, low-income markets continue to exhibit the highest income-price elasticity, especially once the comparison is restricted to identical products. We next include local competition, store size, and product assortment jointly. Table 8 reports the results.

The specification including all controls confirms the previous results. The differential elasticity for Quartile 1 declines relative to the baseline estimates but remains positive and statistically significant in both specifications. With category fixed effects, it falls from 0.023 to 0.018, while with product fixed effects, it declines from 0.020 to 0.017. Thus, even

Table 8: Income Quartile Estimation (Controlled by Store Level Factors).

Dependent Variable:	Log Deflated Price	
Model:	(1)	(2)
Variables		
log(real_income)	0.0651***	0.0428***
	(0.0120)	(0.0064)
Quartile 1	-0.1386***	-0.1334***
	(0.0051)	(0.0002)
Quartile 4	-0.1029***	-0.0332***
	(0.0039)	(0.0006)
Log Number of Cashiers	-0.0001	-0.0023
	(0.0043)	(0.0029)
Share of Store Products	-0.0963***	-0.0535***
	(0.0232)	(0.0136)
Local Store Competition	-0.0005**	-0.0004***
	(0.0002)	(0.0001)
Log Real Income x Q. 1	0.0176***	0.0166***
	(0.0006)	(0.0003)
Log Real Income x Q. 4	0.0111***	0.0034***
	(0.0007)	(0.0002)
Fixed-effects		
Category	Yes	
Time	Yes	Yes
Product		Yes
Fit statistics		
Observations	3,253,651	3,253,651
R^2	0.88429	0.94685
Within R^2	0.02118	0.01842

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. Quartile 1 and Quartile 4 are dummy variables for neighborhoods in the lowest and highest income quartiles, respectively; the omitted group corresponds to the middle-income quartiles. The variables *Log Real Income* $\times$ *Q.* 1 and *Log Real Income* $\times$ *Q.* 4 interact log real income with the first and fourth quartiles. *Log Number of Cashiers* is the logarithm of the number of cashiers in the store and is used as a proxy for store size. *Share of Store Products* is defined as the number of products available in the store divided by the total number of products observed in the sample in that month. *Local Store Competition* is defined as the number of competing stores in the neighborhood, excluding the store itself, divided by neighborhood population density and multiplied by 1,000. The sample combines DGC and Scanntech data; when a store appears in both sources, the DGC observation is retained in the baseline database. Column (1) includes category and time fixed effects, while column (2) includes product and time fixed effects. Standard errors are clustered at the category and market levels.

after simultaneously controlling for local competition, store size, and product assortment, low-income markets remain the most income-elastic.

By contrast, the Quartile 4 differential decreases more markedly, especially in the product fixed-effects specification, where it falls from 0.009 to 0.003. This reinforces the view that the non-linearity in the income-price relationship is concentrated at the bottom of the income distribution rather than reflecting a symmetric pattern across quartiles.

One additional result is worth noting. The share of store products enters negatively and significantly in both specifications, indicating that stores with broader assortments tend to have lower prices even after controlling for local competition and store size. This is consistent with the interpretation that product availability is an important margin behind observed price differences across markets.

Table 9 summarizes how the estimated income-price elasticities evolve as additional controls are introduced. It reports the elasticity for middle-income markets and the differential elasticities for Quartiles 1 and 4 across specifications.

Table 9: Evolution of Income-Price Elasticities across Mechanism Specifications

	Category FE			Product FE		
Specification	Middle	Q1 diff.	Q4 diff.	Middle	Q1 diff.	Q4 diff.
Baseline (Table 3)	0.073	0.023	0.008	0.049	0.020	0.001
+ Local competition (Table 5)	0.062	0.023	0.023	0.041	0.020	0.011
+ Store size (Table 6)	0.067	0.024	0.013	0.045	0.020	0.004
+ Store assortment (Table 7)	0.062	0.018	0.018	0.041	0.017	0.008
+ All controls (Table 8)	0.065	0.018	0.011	0.043	0.017	0.003

Notes: The table reports selected coefficients from quartile-interaction regressions. "Middle" is the coefficient on log real income for the omitted group (Quartiles 2–3). "Q1 diff." and "Q4 diff." correspond to the interaction coefficients between log real income and indicators for Quartile 1 and Quartile 4, respectively. Category FE specifications include category and time fixed effects. Product FE specifications include product and time fixed effects. Each row corresponds to a different specification, progressively adding controls for local competition, store size, and product assortment. The final row includes all controls jointly.

The key pattern is that the Quartile 1 differential declines as controls are introduced, but remains positive and statistically significant across all specifications. By contrast, the Quartile 4 differential is smaller from the outset and becomes close to zero once the comparison is restricted to identical products. This indicates that neither individual controls nor their joint inclusion fully accounts for the stronger income-price elasticity observed in low-income markets. The remaining non-linearity is therefore concentrated at the bottom of the income distribution.

Taken together with the previous subsections, these findings suggest that the income-price relationship reflects multiple layers of heterogeneity. Organizational differences across retailers—particularly the role of chain stores—appear to be the primary source of the non-linearity. Local competitive conditions and store-level characteristics, such as assortment, further contribute to the observed patterns.

The specification, including all controls, confirms this interpretation. The bottom-quartile differential declines relative to the baseline estimates but remains positive and statistically significant, while the top-quartile differential becomes small, especially when the comparison is restricted to identical products. Thus, neither individual control nor their joint inclusion eliminates the main pattern: the non-linearity in the income-price relationship remains concentrated in low-income markets.

Overall, the evidence points to a combination of retail structure, local market conditions, and product availability as the key drivers of the income-price relationship, rather than systematic price differences for identical goods across neighborhoods.

6 Conclusions

This paper examines how retail prices vary with neighborhood income using detailed store–product data from Montevideo. The analysis shows that the positive relationship between income and prices observed in aggregate comparisons declines sharply as the comparison is restricted to more homogeneous goods and retail environments. Once identical products sold by the same retailer are compared, the income–price elasticity becomes economically small.

This finding has a clear implication for how spatial price differences should be interpreted.

The evidence does not support the view that retailers systematically charge different prices for identical goods across neighborhoods. Instead, most of the observed income–price relationship reflects differences in product composition and retail structure across markets. In particular, substitution across products within categories and differences in store types appear to be central margins of adjustment.

At the same time, the income–price relationship is not fully linear. The remaining non-linearity is concentrated at the bottom of the income distribution: low-income neighborhoods exhibit a higher income–price elasticity, while high-income neighborhoods are much closer to the middle-income benchmark. This pattern is primarily driven by chain retailers, suggesting that organizational features of retailing play a central role in shaping price responses to local demand conditions.

Local market conditions and store characteristics contribute to this heterogeneity, but do not fully explain it. Differences in local competition and product assortment account for part of the observed patterns, while store size plays a more limited role. Even after controlling for these margins, part of the heterogeneity remains.

Taken together, these results suggest that spatial price inequality within cities operates mainly through differences in retail environments rather than through differential pricing of comparable goods. This has direct implications for how price dispersion and cost-of-living differences are measured and interpreted. Analyses based on aggregate price indices may overstate the extent to which consumers in different neighborhoods face different prices for the same goods, while understating the role of product availability and retail structure.

More broadly, the findings highlight that understanding spatial inequality requires going beyond prices alone and focusing on the composition of goods, store formats, and local market conditions. From a policy perspective, this suggests that interventions aimed at improving access to retail options—such as entry, competition, or assortment—may be more relevant than those focused on price differences for narrowly defined products.

This paper is descriptive and does not identify a causal effect of income on prices.

Rather, income should be interpreted as a proxy for local demand conditions and neighborhood characteristics. Future work could build on these results by incorporating consumer-level data to study how differences in retail environments translate into differences in prices actually paid and welfare across households. Overall, the evidence suggests that spatial price inequality within cities is shaped mainly by retail environments—product availability, store composition, and local market structure—rather than by systematic price differences for identical goods.

References

Adams, B., and K. R. Williams (2019): “Zone Pricing in Retail Oligopoly,” *American Economic Journal: Microeconomics*, 11(1), 124–156.

Aguiar, M., and E. Hurst (2007): “Life-Cycle Prices and Production,” *American Economic Review*, 97(5), 1533–1559.

Attanasio, O., and C. Frayne (2006): “Do the Poor Pay More?,” Working Paper Series 06-06, The Mario Einaudi Center for International Studies.

Basker, E. (2007): “The Causes and Consequences of Wal-Mart’s Growth,” *Journal of Economic Perspectives*, 21(3), 177–198.

Beatty, T. K. (2010): “Do the Poor Pay More for Food?,” *American Journal of Agricultural Economics*, 92(3), 608–621.

Blow, L., and A. Leicester (2012): “Do the Poor Pay More? An Investigation of British Grocery Purchase Prices,” Working paper series, Institution for Fiscal Studies.

Borraz, F., F. Carozzi, N. González-Pampillón, and L. Zipitría (2024): “Local Retail Prices, Product Variety, and Neighborhood Change,” *American Economic Journal: Economic Policy*, 16(1), 1–33.

Borraz, F., A. Cavallo, R. Rigobon, and L. Zipitría (2016): “Distance and Political Boundaries: Estimating Border Effects under Inequality Constraints,” *International Journal of Finance & Economics*, 21(1), 3–35.

Borraz, F., and L. Zipitría (2012): “Retail Price Setting in Uruguay,” *Economia*, 12(2), 77–109.

——— (2022): “Varieties as a Source of Law of One Price Deviations,” *International Economics*, 172, 1–14.

——— (2024): “Assessing Long-Run Price Convergence in Retailing,” GLO Discussion Paper Series 1424, Global Labor Organization (GLO).

Broda, C., E. Leibtag, and D. E. Weinstein (2009): “The Role of Prices in Measuring the Poor’s Living Standards,” *Journal of Economic Perspectives*, 23(2), 77–97.

Caplovitz, D. (1963): *The Poor Pay More*. The Free Press.

Chung, C., and S. L. Myers (1999): “Do the Poor Pay More for Food? An Analysis of Grocery Store Availability and Food Price Disparities,” *The Journal of Consumer Affairs*, 33(2), 276–296.

Comisión de Promoción y Defensa de la Competencia (2022): “Mercado de Distribución Minorista – Medida Preparatoria,” Discussion Paper Informe No. 23/2022, Ministerio de Economía y Finanzas, República Oriental del Uruguay, Montevideo, Accessed: 2026-04-28.

Daruich, D., and J. Kozlowski (2023): “Macroeconomic Implications of Uniform Pricing,” *American Economic Journal: Macroeconomics*, 15(3), 64–108.

DellaVigna, S., and M. Gentzkow (2019): “Uniform Pricing in U.S. Retail Chains,” *The Quarterly Journal of Economics*, 134(4), 2011–2084.

Dobson, P. W., and M. Waterson (2005): “Chain-Store Pricing Across Local Markets,” *Journal of Economics and Management Strategy*, 14(1), 93–119.

Eizenberg, A., S. Lach, and M. Oren-Yiftach (2021): “Retail Prices in a City,” *American Economic Journal: Economic Policy*, 13(2), 175–206.

Gibson, J., and B. Kim (2013): “Do the urban poor face higher food prices? Evidence from Vietnam,” *Food Policy*, 41(C), 193–203.

Goodman, C. S. (1968): “Do the Poor Pay More?,” *Journal of Marketing*, 32(1), 18–24.

Handbury, J. (2021): "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities," *Econometrica*, 89(6), 2679–2715.

Handbury, J., and D. E. Weinstein (2015): "Goods Prices and Availability in Cities," *Review of Economic Studies*, 82(1), 258–296.

Hitsch, G. J., A. Hortaçsu, and X. Lin (2021): "Prices and promotions in U.S. retail markets," *Quantitative Marketing and Economics (QME)*, 19(3), 289–368.

Jaravel, X. (2021): "Inflation Inequality: Measurement, Causes, and Policy Implications," *Annual Review of Economics*, 13(1), 599–629.

Nakamura, E., and J. Steinsson (2008): "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, 123(4), 1415–1464.

Rao, V. (2000): "Price Heterogeneity And "Real" Inequality: A Case Study Of Prices And Poverty In Rural South India," *Review of Income and Wealth*, 46(2), 201–211.

Stroebel, J., and J. Vavra (2019): "House Prices, Local Demand, and Retail Prices," *Journal of Political Economy*, 127(3), 1391–1436.

A Additional Tables and Figures

Table 10: Descriptive Statistics by Database

Statistic	Average of medians	DGC	Scanntech
Observations		1,828,783	2,279,757
Number of Stores		295	1,038
Neighborhoods		61	61
Total Months Covered		110	110
Median Products per Store	87.50	111.00	64.00
Median Observations per Store	4,104.75	6,791.00	1,418.50
Median Months per Store	65.75	86.00	45.50
Median Cashiers per Store	2.50	4.00	1.00

Source: Authors' calculations based on DGC and Scanntech data.
Notes: The database covers January 2013 to May 2022, excluding January to March 2016, when Scanntech information is unavailable. It includes all grocery and supermarket stores in both databases and is restricted to neighborhoods with available income data.

Table 11: Descriptive Statistics for all Stores by Database

Statistic	Average of medians	DGC	Scanntech
Observations		1,828,783	2,279,757
Number of Stores		295	1,038
Neighborhoods		61	61
Total Months Covered		110	110
Median Products per Store	87.50	111.00	64.00
Median Observations per Store	4,104.75	6,791.00	1,418.50
Median Months per Store	65.75	86.00	45.50
Median Cashiers per Store	2.50	4.00	1.00

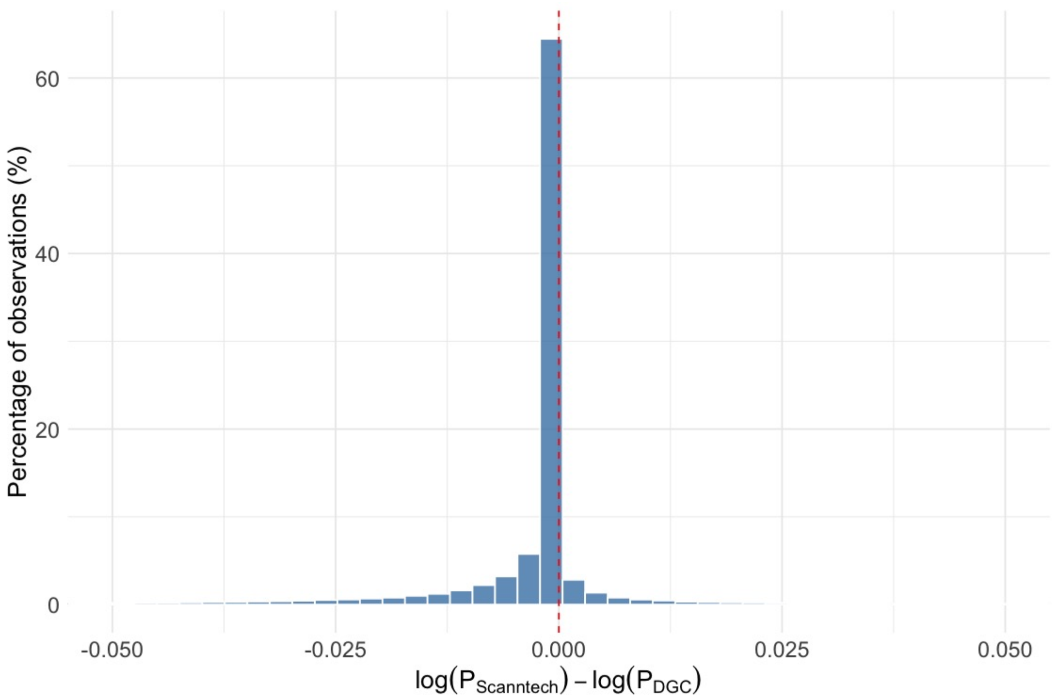
Source: Authors' calculations based on DGC and Scanntech data.
Notes: The database covers January 2013 to May 2022 and excludes January to March 2016, when information from the Scanntech database is unavailable. It includes all grocery and supermarket stores in both databases.

Table 12: Descriptive Statistics by Database for Duplicated Stores Only

Statistic	Average of medians	DGC	Scanntech
Observations		507,807	404,967
Number of Stores		84	95
Neighborhoods		33	35
Total Months Covered		110	110
Median Products per Store	101.00	109.00	93.00
Median Observations per Store	5,247.25	6,790.50	3,704.00
Median Months per Store	67.75	82.50	53.00
Median Cashiers per Store	3.50	3.00	4.00

Source: Authors' calculations based on DGC and Scanntech data.
Notes: The database covers January 2013 to May 2022 and excludes January to March 2016, when information from the Scanntech database is unavailable. It includes stores available in both databases.

Figure 2: Matching Store-Product Price Differences.



Source: Authors' calculations based on DGC and Scanntech data.
Notes: The figure plots the distribution of $\log(P_{Scanntech}) - \log(P_{DGC})$ for matched store-product-month observations. Matches are constructed at the DGC–Scanntech store-pair level and restricted to the common time window for each pair-product combination. The sample includes only observations with non-missing prices in both sources.

Table 13: Price Differences Across Matched Store Pairs

Dependent Variable:	Log Deflated Price		
Model:	(1)	(2)	(3)
Variables			
Scanntech	-0.0017	-0.0017	-0.0017
	(0.0012)	(0.0012)	(0.0012)
Fixed-effects			
Time	Yes	Yes	Yes
Store Pair	Yes	Yes	
Product		Yes	
Store Pair × Product			Yes
Fit statistics			
Observations	639,136	639,136	639,136
R^2	0.14520	0.96488	0.97453
Within R^2	2.48×10^{-6}	6.03×10^{-5}	8.31×10^{-5}

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

Source: Authors' calculations based on DGC and Scanntech data.
Notes: The dependent variable is the log of the deflated price. *Scanntech* is a dummy equal to one for Scanntech observations and zero for DGC observations. The sample includes matched DGC–Scanntech store pairs at the product-month level, restricted to the common time window of each pair-product match. Standard errors are clustered at the store-pair level.

Table 14: Income Elasticity Estimation. Alternative Database.

Dependent Variable:	Log Deflated Price							
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Log Real Income	0.114***	0.067***	0.045***	0.015	-0.198***	-0.114***	-0.022***	-0.452
	(0.001)	(0.011)	(0.006)	(0.027)	(9.4×10^{-5})	(0.003)	(0.0004)	(0.270)
Scanntech Database	0.004	0.016	0.012		0.004	0.016	0.012	
	(0.042)	(0.014)	(0.009)		(0.042)	(0.017)	(0.010)	
(Log Real Income)^2					0.018***	0.010***	0.004***	0.026*
					(0.0002)	(0.0009)	(0.0004)	(0.015)
Fixed-effects								
Time	Yes	Yes	Yes		Yes	Yes	Yes	
Category		Yes				Yes		
Product			Yes	Yes			Yes	Yes
Store				Yes				Yes
Fit statistics								
Observations	3,150,811	3,150,811	3,150,811	3,150,811	3,150,811	3,150,811	3,150,811	3,150,811
R^2	0.00588	0.89366	0.94996	0.95180	0.00592	0.89367	0.94997	0.95185
Within R^2	0.00418	0.01391	0.01335	0.00066	0.00422	0.01403	0.01338	0.00162

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Source: Authors' calculations based on DGC and Scanntech data. *Notes:* The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. The dummy *Scanntech Database* equals one for observations from the Scanntech database and zero for observations from the DGC database. Columns (1)–(4) include income only in linear form, while columns (5)–(8) add a quadratic income term. The specifications progressively introduce fixed effects for time, category, product, and store, as indicated in the table. Standard errors are clustered at the category and market levels.

Table 15: Chain Store Robustness: Including Chain Fixed Effects

Dependent Variable:	Log Deflated Price	
is.chain	Chains	
Model:	(1)	(2)
Variables		
log(real_income)	0.001	0.002
	(0.005)	(0.003)
Quartile 1	-0.163***	-0.142***
	(0.054)	(0.043)
Quartile 4	-0.065***	-0.010**
	(0.011)	(0.005)
Log Real Income x Q. 1	0.020***	0.018***
	(0.006)	(0.005)
Log Real Income x Q. 4	0.007***	0.001**
	(0.001)	(0.0005)
Fixed-effects		
Category	Yes	
Time	Yes	Yes
Chain	Yes	Yes
Product		Yes
Fit statistics		
Observations	1,482,915	1,482,915
R^2	0.87871	0.94706
Within R^2	9.91×10^{-5}	0.00011

Clustered (Category & Market) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The dependent variable is the log of the deflated price. Log real income is measured at the neighborhood-month level. Quartile 1 and Quartile 4 are dummy variables for neighborhoods in the lowest and highest income quartiles, respectively; the omitted group corresponds to the middle-income quartiles. The variables $LogRealIncome \times Q.1$ and $LogRealIncome \times Q.4$ interact log real income with the first and fourth quartiles. The sample is restricted to chain stores only. Column (1) includes category, time, and chain fixed effects, while column (2) includes product, time, and chain fixed effects. Standard errors are clustered at the category and market levels.

B Construction of the Price Database

This appendix describes the construction of the price database used in the empirical analysis. The final dataset combines two original sources: the price records collected by the *Dirección General de Comercio* (DGC) and the scanner data provided by the private firm Scanntech, based on store-reported sales. The construction procedure had three objectives: to build a comparable monthly panel at the store–product level in each source, to harmonize products and stores across sources, and to identify overlapping establishments before combining both datasets into a single analytical file.

B.1 DGC data

Stores have to report their daily prices to the DGC each month if they meet either a size criterion —have at least three cashiers—or have some stores within the same commercial name—at least four—, and sell a minimum share of products listed by the DGC. The DGC source is originally recorded at the daily level. The raw files are first cleaned in order to produce a consistent daily store–product panel. Exact duplicates at the store–product–date level are removed. When multiple observations remain for the same store, product, and day, identical prices are treated as redundant duplicates, whereas conflicting daily prices are discarded, since the true price cannot be recovered unambiguously from the raw source.

The cleaned daily data are then used to construct a monthly price series. Before aggregation, outliers are removed using a proportional rule defined at the product–month level. For each observation, the median price of the same product in the same month is computed, and the observation is retained only if it lies within the interval

$$\left[\frac{1}{3}\widetilde{p}_{jm},\, 3\widetilde{p}_{jm}\right],$$

where $\widetilde{p}_{jm}$ denotes the median price of product j in month m. This rule removes implausibly low and high prices while preserving comparability across products with different price levels. The removed prices account for less than 0.1% of the observations.

The monthly DGC series used in the analytical database corresponds to the average monthly price at the store–product level. This aggregation is the natural counterpart to the Scanntech monthly average price. We restrict the sample, which begins in March 2007, to the period shared with Scanntech. We discard products that are either not homogeneous, such as various cuts of meat and brandless bread, or sold unpackaged by stores, such as chicken, sausage, and ham. These products are identified using the official DGC product list and removed from the sample. Also, we removed the mayonnaise "Fanacoa" because it is not available in Scanntech. Store-level information is then added from a cleaned DGC establishments file, which contributes neighborhood, chain affiliation, and number of cashiers. After this merge, one duplicate establishment is removed, and stores with very limited temporal coverage are excluded, defined as having fewer than one observation per period. Lastly, because the Scanntech database lacks data for January to March 2016, those months were also excluded.

B.2 Scanntech data

The Scanntech source is already organized at a monthly frequency, but it is originally distributed across several files. These files are first appended into a single database, the overlap around 2017 is resolved, and variable names are standardized. The sample is then restricted to Montevideo by normalizing locality labels and dropping observations outside the department.

The next stage is to construct a valid store-level file. Store names and addresses are standardized, and outlets not comparable to the target supermarket and self-service segment are excluded. This includes, among others, butcher shops, pharmacies, kiosks, bakeries, greengrocers, gas stations, and other specialized formats. Stores with missing or invalid addresses are also dropped, since address information is required both for cleaning and for the subsequent cross-source matching.

Chain affiliation is then completed or standardized using the point-of-sale name. As in the DGC database, stores with very limited temporal coverage are removed, and a harmonized store identifier is assigned from a pre-existing store ID file. Neighborhood information is constructed by combining a spatial assignment based on coordinates and the Montevideo shapefile with manual review based on addresses. Cases that remain unresolved after the automatic procedures are completed are handled manually.

Product harmonization is designed to reproduce the DGC basket as closely as possible. The list of unique Scanntech products is linked to DGC product identifiers using barcode strings. As in the DGC data, product codes beginning with 9 are excluded. Matching errors are corrected manually when necessary, and additional filters are applied to eliminate presentations that are not exact matches even if they belong to the same product family. The resulting product file assigns each Scanntech item a DGC product identifier and a category.

The final Scanntech price panel is obtained by merging the intermediate monthly database with the cleaned product file and the validated store file. The database keeps one observation per store, product, and month, drops observations with missing neighborhood information, and constructs the same time index used for DGC. Exact duplicates at the store–product–time level are then removed. Outliers in Scanntech are treated using the same proportional rule as in DGC.

C Income Database

This appendix describes the construction of the neighborhood-level income database used in the empirical analysis. The database is built from household survey information provided by the *Instituto Nacional de Estadística* (INE) and covers Montevideo neighborhoods at the monthly level.

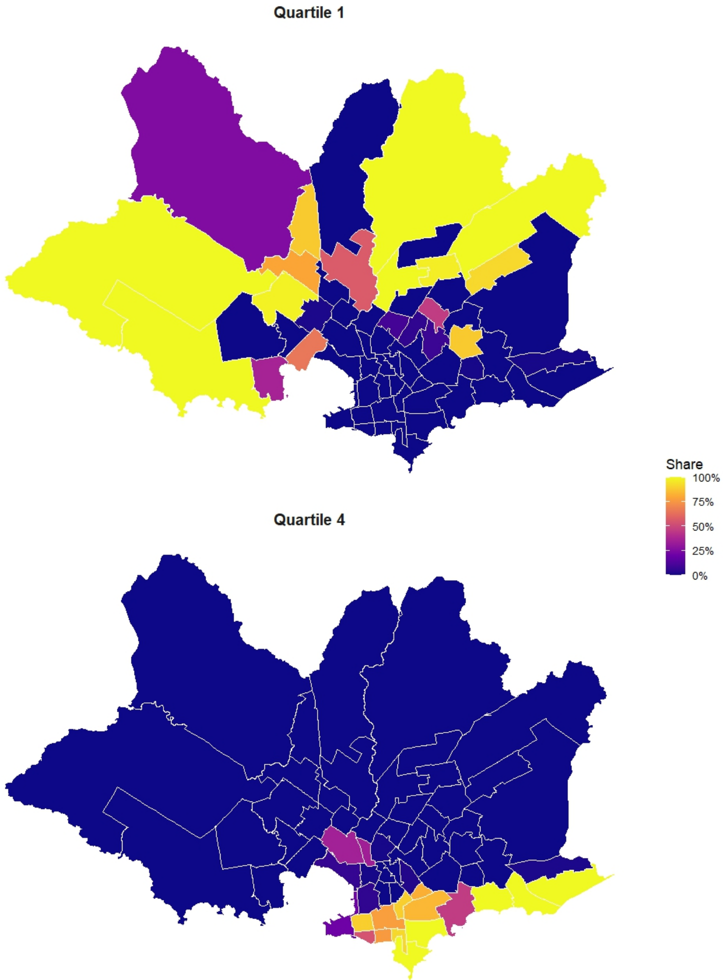
Neighborhood-level income is subject to sampling variability because some neighborhood-month cells contain few household observations. This generates noise in raw monthly income measures, particularly in smaller neighborhoods. The procedure described below is designed to recover a smooth and economically meaningful measure of local purchasing power while minimizing the influence of extreme values and missing observations. This construction aims to approximate local purchasing power faced by consumers, which is the relevant dimension for analyzing price differences across neighborhoods.

The final income database includes 49 Montevideo neighborhoods: Aguada, Aires Puros, Atahualpa, Barrio Sur, Belvedere, Brazo Oriental, Buceo, Capurro–Bella Vista, Carrasco, Carrasco Norte, Casabó–Pajas Blancas, Casavalle, Castro–P. Castellanos, Centro, Cerrito, Cerro, Ciudad Vieja, Colón, Conciliación, Cordón, Ituzaingó, La Blanqueada, La Comercial, La Teja, Larrañaga, Lezica–Melilla, Malvín, Malvín Norte, Manga–Toledo Chico, Maroñas–Parque Guaraní, Mercado Modelo–Bolívar, Nuevo París, Palermo, Parque Batlle–Villa Dolores, Parque Rodó, Paso de la Arena, Peñarol–Lavalleja, Piedras Blancas, Pocitos, Prado–Nueva Savona, Punta Carretas, Punta Gorda, Punta Rieles–Bella Italia, Reducto, Sayago, Tres Cruces, Unión, Villa Española, and Villa García–Manga Rural.

In the INE classification, Colón is split into two areas: Colón Centro y Noroeste and Colón Sureste–Abayubá. We collapse these two areas into a single neighborhood, Colón, to maintain consistency with the neighborhood definitions used in the price database. The following neighborhoods are excluded due to insufficient observations to construct reliable monthly income measures: Bañados de Carrasco, Flor de Maroñas, Jacinto Vera, Jardines del Hipódromo, La Figurita, La Paloma–Tomkinson, Las Acacias, Las Canteras, Manga, Paso de las Duranas, Tres Ombúes–Victoria, and Villa Muñoz–Retiro.

Figure 3 shows, for each neighborhood, the share of periods in which it is classified in Quartile 1 and Quartile 4, highlighting the spatial persistence of income differences across neighborhoods.

Figure 3: Spatial Distribution of Neighborhood Income Quartiles over Time.



Notes: The figure reports the share of periods in which each neighborhood is classified into the lowest and highest income quartiles. The upper panel shows the share of neighborhood-time observations in which the neighborhood belongs to Quartile 1, while the lower panel shows the corresponding share for Quartile 4. The underlying data are collapsed to one observation per neighborhood and time period, and classification is based on the neighborhood's income quartile in each period. Darker shades indicate a higher share of periods spent in the corresponding quartile.

C.1 Income Imputation Procedure

First, we construct neighborhood income using household survey information from INE. Income is measured net of taxes, consistent with the definition used in the household survey, and captures the disposable income available to households for consumption. For each neighborhood, monthly income is computed as a six-month rolling average of current and lagged income. Let y_{it} denote the resulting six-month average income for neighborhood i in month t:

$$y_{it} = \frac{1}{6} \sum_{s=t-5}^{t} y_{is}^{raw},$$

where y_{is}^{raw} is the average income observed in neighborhood i in month s. This rolling average is computed prior to the outlier detection and interpolation steps described below.

Because some neighborhood-month cells have missing values or display sharp fluctuations due to limited household observations, we apply a cleaning procedure to the resulting monthly neighborhood income series. Let i index neighborhoods and t months. The variable of interest is neighborhood mean income, denoted by $income_{it}$. All computations are performed after sorting the data by neighborhood and month.

The procedure consists of two steps: (i) identification and treatment of outliers, and (ii) interpolation of missing values.

Outlier detection

Outliers were identified using within-neighborhood monthly changes in log income. For each neighborhood i, we computed

$$\Delta \log(y_{it}) = \log(y_{it}) - \log(y_{i,t-1}). \tag{3}$$

We then calculated the neighborhood-specific median absolute deviation (MAD) of this monthly change series,

$$MAD_i\Big(\Delta \log(y_{it}\Big) = \text{median}_t \left(\Big|\Delta \log(y_{it}) - \text{median}_t\Big(\Delta \log(y_{it})\Big)\Big|\right). \tag{4}$$

An observation was classified as a soft outlier if its absolute monthly change exceeded the larger of a fixed threshold and three times the neighborhood-specific MAD:

$$|\Delta \log(y_{it})| > \max\Big\{0.10,\ 3 \times MAD_i\Big(\Delta \log(y)\Big)\Big\}. \tag{5}$$

An observation was classified as a hard outlier if

$$|\Delta \log(y_{it})| > 0.20. \tag{6}$$

In the cleaning stage, hard outliers were recoded as missing values prior to interpolation. In addition, zero values were also treated as missing.

Interpolation

After removing hard outliers and recoding zeros as missing values, the income series was completed in three steps.

First, we constructed a balanced neighborhood-month panel by generating the full sequence of months for each neighborhood. This ensured that all missing months were explicitly represented.

Second, gaps of one or two consecutive months were filled by linear interpolation in logarithms. Let $\mathcal{I}[\cdot]$ denote linear interpolation across adjacent non-missing observations. For short gaps, imputed income was defined as

$$\widetilde{y}_{it} = \exp\Big(\mathcal{I}\Big[\log(y_{it})\Big]\Big)\,. \tag{7}$$

Third, any remaining missing observations were imputed using Kalman smoothing based on a univariate structural time-series model estimated separately for each neighborhood. The model was applied to the log income series and used to recover values not filled by short-gap interpolation.

As a final fallback, residual missing observations were imputed using the cross-sectional structure of the city. For each month t, we computed the citywide median income across neighborhoods, denoted by $income_t^{city}$. For each neighborhood i, we then estimated a stable neighborhood-specific share,

$$s_i = \text{median}_t\left(\frac{y_{it}}{y_t^{city}}\right), \tag{8}$$

using months with valid observations. Residual missing values were then set equal to

$$\widetilde{y}_{it} = s_i \, y_t^{city}. \tag{9}$$

Overall, this procedure produces a stable and continuous measure of neighborhood income that preserves meaningful temporal variation while limiting the influence of sampling noise. This is particularly important for our analysis, which relies on cross-neighborhood differences in purchasing power to identify price-income relationships.

D Disclosure

AI Usage Disclosure During the preparation of this work, the authors used OpenAI services for the following tasks:

- Writing refinement.
- Title brainstorming.
- Detecting writing errors and ensuring consistent and concise language usage.
- Coding support to improve the efficiency of computing and to improve/refine visualization techniques.
- Help with Latex issues and formatting, particularly in creating and designing tables and figures.

After using this OpenAI service, the authors reviewed and edited the content as needed and take full responsibility for the content of the paper.