Are Poor Cities Cheap for Everyone?

Non-Homotheticity and the Cost of Living Across U.S. Cities*

Jessie Handbury†

April 20, 2021

Abstract

This paper shows that the products and prices offered in markets are correlated with local income-specific tastes. To quantify the welfare impact of this variation, I calculate local price indexes micro-founded by a model of non-homothetic demand over thousands of grocery products. These indexes reveal large differences in how wealthy and poor households perceive the choice sets available in wealthy and poor cities. Relative to low-income households, high-income households enjoy 40 percent higher utility per dollar expenditure in wealthy cities, relative to poor cities. Similar patterns are observed across stores in different neighborhoods. Most of this variation is explained by differences in the product assortment offered, rather than the relative prices charged, by chains that operate in different markets.

Keywords: Non-homotheticity, price index, variety, cost of living.

*This paper was the main chapter of my Ph.D. dissertation. I am grateful to my advisors David Weinstein and Kate Ho for invaluable advice. I thank Songyuan Ding, Yue Cao, and Serena Xu for outstanding excellent research assistance. I also thank David Bieri, Chris Conlon, Don Davis, Jonathan Dingel, Gilles Duranton, Galina Hale, Jean-Francois Houde, Amit Khandelwal, Joan Monras, Aviv Nevo, Eleonora Patacchini, Mike Riordan, Molly Schnell, Katja Seim, Ina Simonovska, Todd Sinai, Daniel Sturm, Eric Verhoogen, Jonathan Vogel, and Maisy Wong and several presentation attendees for valuable discussions and comments. I gratefully acknowledge financial support from the Program of Economic Research at Columbia and the Research Sponsors' Program of the Wharton Zell/Lurie Real Estate Center. Researcher's own analyses are calculated (or derived) based in part on data from the Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. All errors are my own.

†The Wharton School, University of Pennsylvania, 431 Vance Hall, 3733 Spruce St, Philadelphia PA 19094-6302, handbury@wharton.upenn.edu.

1 Introduction

It is well known that prices and product variety vary systematically across space: high-end goods are more available in rich neighborhoods than poor ones. Yet the cost-of-living indexes that economists employ to account for these spatial price differences aggregate prices using the same expenditure weights for all consumers, implicitly assuming that tastes do not vary with income.[1] Under this assumption, a high-income Washington D.C. resident would be indifferent between the set of goods available in their local stores and the set available in a city with less than half the per capita income, like Detroit. In reality, preferences are non-homothetic (see, *e.g.*, Deaton and Muellbauer (1980) and Bils and Klenow (2001)). This paper is the first to study the implications of non-homotheticity for spatial price indexes.

I first document how availability and prices of grocery products varies with local income across U.S. cities as well as across neighborhoods within these cities. To measure the implications of these spatial availability and pricing patterns for the welfare of consumers at different income levels, I next develop a model of non-homothetic demand. I estimate the model with a combination of data describing the aggregate sales of different products in a sample of stores across the U.S. and the purchases of individual households in those stores. I use the estimated model to construct price indexes that summarize how households at different income levels value the prices and products available to them in different geographic markets. Finally, I characterize how and why the price level varies across cities and neighborhoods in the U.S. differently for consumers at different income levels. This analysis yields three sets of novel results.

First, stores favor high-income consumers more in wealthy locations than in poor ones through both their pricing and product offerings. Stores in wealthier cities offer products representing a greater share of the high-income consumption bundle than the low-income consumption bundle. Stores in wealthier cities also charge relatively less for the high-income consumption bundle than the low-income one, conditional on availability. The same patterns are observed across stores in different neighborhoods of the same city.

Second, these differences in availability and pricing matter for consumers. Income-specific spatial price indexes reveal large differences in how high- and low-income households perceive the prices and variety available in different U.S. cities. Once you account for income-specific tastes, markets that are relatively expensive for poor households can be instead relatively cheap for the wealthy. For example, a low-income household earning \$25,000 a year faces 9 percent

[1] Albouy (2009) and Moretti (2013), for example, use the ACCRA indexes of intra-national price variation, while Deaton (2010) and Almas (2012) calculate homothetic indexes of international price variation. The importance of recognizing the these non-homotheticities in regional price indexes was recognized over 50 years ago in Snyder (1956). Related work has considered the impact of non-homotheticity in demand for purchasing power parity deflators (Deaton and Dupriez, 2011a; Li, 2021) and real income inequality (Albouy et al., 2016).

higher grocery costs in Bridgeport, CT, with per capita income \$50,000, relative to Flint, MI, with per capita income below \$25,000. But the same is not true for high-income households earning \$200,000 a year whose grocery costs are 19 percent lower in Bridgeport than in Flint.

Third, I show that the differences in relative grocery costs across cities are driven more by cross-city variation in product variety than by variation in prices. Higher income households find groceries cheaper in wealthier cities primarily because more varieties of the high-quality products that high-income consumers prefer to consume are available in these locations. These high-quality products are sold at lower unit prices relative to low-quality products in wealthy cities, but these price differences only explain a small portion of the gap between the grocery costs perceived by high- and low-income households across wealthy and poor cities. This result points towards a second short-coming of conventional price indexes, which compare only the prices of common goods, and not variety differences, across locations.[2] Even if they are non-homothetic, price indexes that do not account for differences in product availability will fail to capture any of the true cost-of-living differences for wealthy, relative to poor, consumers.

I also study how store-level price indexes vary across and within cities. I find that higher income households face relatively lower price indexes in stores located in higher income neighborhoods, even within the same CBSA. In fact, the cross-CBSA variation in income-specific price indexes is strongest between stores located in above median neighborhoods within each CBSA. Thus, within-city sorting can maximize a wealthy consumer's variety gains from living in a wealthy city, and mitigate the relative losses for a poor consumer. I finally use the store-level indexes to better understand why variety varies across and within cities. Here I find that the variation in variety offerings across CBSAs and neighborhoods is entirely driven by variation in the local mix of retail chains. There is no systematic variation in the price indexes high- and low-income households face across stores belonging to the same retail chain.

The main methodological challenge I overcome in this paper is to summarize the costs that consumers face across multiple differentiated product categories in a way that parsimoniously accounts for the non-homothetic tastes demonstrated in household behavior. To do this, I build income-specific price indexes. A major reason why existing regional price indexes do not take non-homotheticities into account is that the single-sector models used to identify non-homotheticities in micro studies do not lend themselves to aggregation. I nest a variant of these micro models, from the log-logit/constant elasticity of substitution (CES) family, in a Cobb-Douglas superstructure to model non-homothetic preferences across differentiated products in many sectors. Log-logit sub-utility functions govern how idiosyncratic consumers allocate ex-

[2]Handbury and Weinstein (2014) find a huge amount of variation in availability of grocery varieties across U.S. cities and show that conventional price indexes underestimate the correlation between city size and the grocery price level, for a homothetic representative consumer, by about a third. Variety differences play a much larger role here, explaining all of the positive correlation between city income and the differences in the grocery price levels faced by wealthy, relative to poor, consumers.

penditures between products within product categories, while Cobb-Douglas utility governs the substitutability of products across different categories. The key feature of this structure is that it can be aggregated in such a way that one could also express aggregate product demand as if it had been derived from a representative (non-homothetic) household.[3] This provides a way to bridge the gap between the micro data that I use to identify parameters and an aggregate non-homothetic price index that can be used to compare price levels across locations.

The model nests two forms of non-homotheticity and is structured in a way that enables me to test for their relative importance in explaining the differences between the purchases of high- and low-income consumers. The elasticity of demand with respect to price and product quality depends on the consumer's expenditure on a composite of non-grocery products which I assume to be normal. The intuition here is that, if high-income households spend more on cars, schooling, and housing, for example, then they have a greater willingness to pay for their own ideal product variety or for products that are ranked as high quality by all consumers. These are the most common ways in which international economists hypothesize that non-homotheticities might matter (Hummels and Lugovskyy (2009), Simonovska (2015), Fajgelbaum et al. (2011), and Faber (2014)).[4] Where previous papers have verified each of these channels of non-homotheticity independently, this is the first to test their empirical relevance concurrently and to assess their relative importance in explaining consumer behavior. My results demonstrate the salience of non-homothetic demand for quality in U.S. grocery consumption. I compare three different models of non-homotheticity: a specification in which the taste for quality rises with income, a specification in which high-income households are less price sensitive, and a specification in which both factors play a role. I find that the specification that allows for non-homothetic demand for quality alone explains the differences between the purchases of rich and poor households most parsimoniously.[5]

The main contribution of this paper is to provide the first direct evidence of income-specific tastes for local consumption amenities. A recent urban economics literature hypothesizes that these tastes may help explain spatial disparities in income and skill observed across U.S. cities: high-skill, high-income workers co-locate because they enjoy more utility from certain endogenous local amenities than low-skill, low-income consumers (see, *e.g.*, Glaeser et al. (2001), Diamond (2016) and Couture and Handbury (2020)). Previous empirical support of this theory

[3]The origins of this result are Anderson et al. (1987), whose proof is extended to models that account for product quality in Verhoogen (2008). This link has also been explored in Hortaçsu and Joo (2015) who present a generalized version of the demand system developed here that allows for tastes for product quality to vary with both observed and unobserved consumer attributes.

[4]There are other reasons that demand may vary with income, related to demand for variety (Li (2021)) and shopping behavior (Aguiar and Hurst (2005)). These do not appear to be the primary factors driving differences in the purchases of high- and low-income households in this dataset and are, therefore, not included in the model.

[5]Faber and Fally (2017) estimate the same demand system non-parametrically using only the household-level data and also find that the differences in price elasticities across income quintiles are small relative to the cross-quintile differences in the elasticities of demand for quality.

relies on spatial equilibrium models that assume people are perfectly mobile, inferring changes in skill-biased amenities as those which reconcile changes in housing price and wage data with observed changes in the skill composition of U.S. cities (Diamond (2016), Black et al. (2009)). I instead measure these skill-specific amenities directly, providing cross-sectional evidence that non-housing price indexes are correlated with local incomes in such a way that might encourage further skill-biased agglomeration.

In particular, I show that product variety is skewed towards the income-specific tastes of local consumers. This result is consistent with the theory that, in markets with increasing returns and demand heterogeneity, differentiated product firms cater to local tastes generating "preference externalities" or "home market effects." Fajgelbaum et al. (2011), for example, show theoretically that high-income consumers with non-homothetic preferences enjoy greater consumption utility when living in high-income countries. Like Waldfogel (2003), I provide evidence suggesting that the mechanism behind these effects is local distributors catering to local tastes. My main contribution here, however, is to demonstrate the economic significance of these externalities by measuring their impact on consumer costs. My results showing that these preference externalities are mediated by chain-level pricing and product assortment decisions corroborate a growing literature on these decisions (DellaVigna and Gentzkow (2019); Hitsch et al. (2019); Adams and Williams (2019)) and the role that they play in generating cross-city variation in aggregate variety (Hottman (2014)).[6,7]

These results have mixed implications for the question of how to account for cost-of-living differences across locations when measuring welfare. Standard homothetic price indexes implicitly ignore that households with different incomes have different tastes and, therefore, may perceive these relative costs differently. I find that these cost differences are large in the context of non-durable goods. If similar group-specific externalities are at play in other non-tradable sectors (such as housing, non-tradable services, and durables), it may be necessary to account for income-specific tastes when measuring relative real incomes and expenditures of households at opposite ends of the income distribution. Such adjustments may, for example, have implications for the recent findings on how ignoring intra-national price variation biases measures

[6]The observed distribution of product availability is also consistent with a comparative advantage story and my analysis does not identify this story from the preference externalities. Dingel (2016) shows that the specialization of high-income counties in exporting high-quality products is explained as much by home-market demand as by differences in factor usage and endowments.

[7]Complementary work finds variation in inflation across income groups. The BLS has a long tradition of using confidential survey data to construct inflation indexes that use income-specific expenditure weights (see, *e.g.*, Snyder (1961); Kokoski (1987); Jorgenson et al. (1989); Garner et al. (1996); Cage et al. (2002)). More recent papers apply a method developed by Broda and Romalis (2009) to calculate income-specific exact price indexes for the U.S. with the same household purchase data used here (Argente and Lee, 2016; Jaravel, 2018). On the structural side, Albouy et al. (2016) quantify a model of non-homothetic housing demand to show that the poor have been disproportionately impacted by rising relative rents in the U.S., and Atkin et al. (2020) use an AIDS model to calculate aggregate income-specific inflation rates for Indian households.

of real income inequality (Moretti, 2013; Albouy et al., 2016) and the geographic distribution of real tax expenditures in the U.S. (Albouy, 2009). Finally, these results suggest that it may also be worth revisiting whether to use homothetic price indexes to account for location-specific costs when calculating poverty thresholds or entitlement payments, as is undertaken in Deaton and Dupriez (2011b).

2 Data

The analysis in this paper is based on detailed store sales and household purchase data, provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. I use the store sales data to infer the set of products and prices available in U.S. cities and the household purchase data to identify how consumers at different income levels value these products and prices. These two Nielsen datasets are available from 2006 onward. I analyze data from a single year, 2012, during which I assume there is no intertemporal variation in the product set and tastes. I complement the 2012 Nielsen data with 5-year 2010-2014 average of tract- and CBSA-level population and income data from the American Community Survey (ACS accessed via the NHGIS, Manson et al. (2018)) to measure how prices and product availability co-vary with local wealth across cities and neighborhoods. In what follows, I describe the structure of each Nielsen dataset and the key variables I draw from them. Further details are available in Appendix A.

The Nielsen store-level (RMS) data contains a panel of weekly sales and quantities by Universal Product Code (UPC) collected by point-of-sale systems in over 30,000 participating retailers across the U.S., along with the county in which each store is located. I complement the RMS data with the Nielsen household-level (HMS) data, which contains information on all bar-coded product purchases made by a panel of over 100,000 households in markets across the United States. Each household in this sample was provided with a bar-code scanner and instructed to collect information such as the UPC, the value and quantity, the date, and the name, location, and type of store for every purchase they made. Nielsen also surveys each household to collect information on, among other things, income, household size, and residential 5-digit zip code.

The RMS data is collected in an automated process so it is less prone to measurement error than the HMS household survey data. As such, the RMS data is better-suited for the construction of non-linear sales share moments I use to identify price elasticity and quality parameters common to all households. The HMS data, meanwhile, provides a detailed picture of the products selected by households at different income levels in the same store and is useful for documenting differences in purchases by income level, controlling for their choice set, and estimating the parameters that generate these differences in the model.

The HMS data also allows me to obtain a more precise estimate of household income in the neighborhood surrounding each store. I measure the income distribution in a store's vicinity with a distance-weighted average of the income distributions observed in the Census tracts within 30km of the centroid of the modal residential zip code of Nielsen panelists that report shopping at that store over all available years (2006 through 2017).

The demand estimation procedure employs only those household-level purchases that are made in RMS retailers. Along with the data cleaning steps outlined in Appendix A.1, this limits the sample of purchases employed for estimation to around 10 percent of the expenditures in the raw data.[8]

Product Definitions

Nielsen categorizes UPCs into "modules." Within each module, I aggregate UPCs into a classification that I call a "product." A product is defined as the set of UPCs within a module with the same brand. For example, in the module "SOFT DRINKS - CARBONATED", there are 104 UPCs that refer to drinks sold under the brand "COCA-COLA R" (R stands for regular, as opposed to diet). These UPCs belong to the same product.[9]

Table 1 shows how UPCs are distributed across products and modules in the sample used to estimate demand. This sample has been cleaned in various ways. To ensure that differences in container sizes or multi-packs do not mechanically generate spurious differences in prices in my sample, I define prices on a per unit basis throughout the paper, using the modal unit definition for each module. I limit my attention to products whose container size is expressed in the modal units for their module and exclude modules whose modal container size is either not expressed in meaningful units (*e.g.*, counts instead of weights or volume) or in the same units for at least 75% of UPCs.[10] To avoid differences in product quality that could be correlated with store amenities or neighborhood income, I exclude random weight items.[11] To control for data recording errors, I drop any store-month in which I observe a UPC sold at a unit price greater than three times or less than a third of the median unit price paid per unit of any UPC within the same product or module categorization. For computational reasons, I put products whose

[8]The similarity of the headline results here with those in earlier drafts that used only, but all of, the household-level purchase data for estimation indicates that this sample restriction does not introduce significant bias.

[9]The analysis abstracts from other product characteristics, such as container, flavor, size, and whether the product was sold in a multi-pack or not. Differentiating between products along these dimensions leads to many products with sales shares too low to allow for the matrix inversions required in the estimation procedure.

[10]Approximately one quarter of modules do not satisfy this restriction. Within the modules that are included, products whose container size is not expressed in the modal units for the module represent 1.3% of store sales in the RMS data.

[11]The quality of random weight items, such as fruit, vegetables, and deli meats, varies over time as the produce loses its freshness and it is likely that stores set prices to reflect this. This potential inter-temporal correlation between their unobserved quality of random weight products and their prices would introduce biases in the price elasticities estimated below.

average positive sales shares across CBSA-month markets fall below the 60th percentile into an outside product and drop sales from any markets that sell less than two non-outside products. Finally, for identification purposes, I limit my attention to modules that have some overlap between the product-store-month RMS store sales data and the HMS household purchase data and to products that are sold in 5 or more of the remaining markets. The cleaned data contains approximately 260,000 UPCs categorized into approximately 37,000 products across over 700 product modules. Almost two thirds of these products are purchased by households in the HMS data. The median numbers of products and UPCs per module are 39 and 118, respectively.

Table 1: Summary Statistics for the Nielsen Data Used in Estimation

Data:	RMS (Store)							HMS (HH)
	Total	Count Per Module			Count Per Product			Total
	Count	Min	Median	Max	Min	Median	Max	Count
Modules	708	-	-	-	-	-	-	708
Products	37,284	2	39	766	-	-	-	24,987
UPCs	266,277	2	118	8,546	1	6	1,347	139,443

Notes: This table shows the distribution of UPCs across product and module categories in the Nielsen RMS store sales and HMS household purchase data used for estimation. This estimation sample has been cleaned from the raw Nielsen data as described in Section 2 of the paper. A product is defined as the set of UPCs within a module with the same brand. The table does not include the “outside” product (into which 60 percent of products are allocated, in the base specification).

The utility function presented below assumes that, conditional on price, consumers do not differentiate between UPCs in the same product. The assumption might be violated in cases where different UPCs that I have defined to be the same product are differentiated by their packaging or flavor. To check the extent to which consumers differentiate between UPCs within product categories, I compared the coefficient of variation for the average unit price paid for each UPC with the coefficient of variation for the average unit price paid for the set of UPCs with the same product categorization. The median coefficient of variation of unit values across UPCs in a given module is 0.51, only slightly higher than the median coefficient of variation of unit values across products in a given module at 0.50, and the two statistics are highly correlated across modules ($\rho = 0.96$). This indicates that there is little variation in the prices charged for UPCs within the same product.

Household Income

The Nielsen HMS data is uniquely suited for estimating how consumers at different income levels value products because it links detailed information on household purchases to information on their reported annual income and demographics. Nielsen classifies households into 16 brackets of reported income. For my analysis, I exclude households with reported incomes below \$11,000 and/or missing demographic data. I convert household income to a continuous variable equal to the mid-point of the income range represented by their Nielsen income category and an income of \$150,000 to the households in the "above \$100,000" income category. I then adjust income for household size using a square-root equivalence scale.[12]

Nielsen under-samples low-income households and, to a lesser degree, high-income households (see Appendix Figure A.2), but has positive weight of households at most income levels – up to the top-code – which, combined with functional form assumptions, allows for the calculation of price indexes at all points along the income distribution.

City-Level Product and Price Availability

I infer the products and prices available in CBSAs in 2012 with those that I observe in the sales of local outlets of Nielsen participating retailers in that year. Not all stores participate in the RMS sample, so I likely observe only a sub-set of the products available in each city. This sample might not be representative, so the product availability and prices in the raw data will be subject to biases related to the number and type of stores sampled in each city.[13] To deal with these potential biases, I infer CBSA-level product availability and pricing using the sales of randomly-selected sub-samples of stores from each city. For the main analysis, I use products and unit prices represented in the sales of 50 randomly-selected stores, limiting my attention to 125 cities with 50 or more retailers in the RMS sample.[14]

In the analysis comparing pricing and product availability across stores, I limit attention to grocery stores (listed in the Nielsen data as in the "food" channel), dropping mass merchandisers, drug, and convenience stores, which may exhibit different relative pricing and availability patterns.

[12] This simple rule of thumb has been employed by the OECD Income Distribution Database (IDD) since 2012 (http://www.oecd.org/els/soc/IDD-ToR.pdf). The bulk of the resulting distribution of size-adjusted income for the households considered in the analysis is between \$10,000 and \$80,000, which seems reasonable given that the per capita incomes of the cities represented in the sample ranges from approximately \$30,000 to \$60,000.

[13] This data limitation is common to all work that builds spatial price indexes from micro data. A key concern here is sampling bias towards stores in higher-income neighborhoods. Appendix Figure A.2 shows that the Nielsen participating retailer sample is over-weighted towards stores in higher-income neighborhoods, relative to the distribution of grocery stores in the County Business Patterns zip-level data, but only to a small degree.

[14] Appendix A.4 lists the population and total number of sample stores for the 125 cities considered in this analysis. Sampling stores in proportion to the total number of stores or the density of stores in each CBSA yields more pronounced differences in product availability between high- and low-income cities than sampling a fixed count of stores from each CBSA. Appendix B shows that the skew in the product variety available in high-income cities towards products favored by high-income households is three times as large when product variety in each CBSA is inferred using the sales of a proportional number of stores instead of a fixed count of stores.

3 Stylized Facts

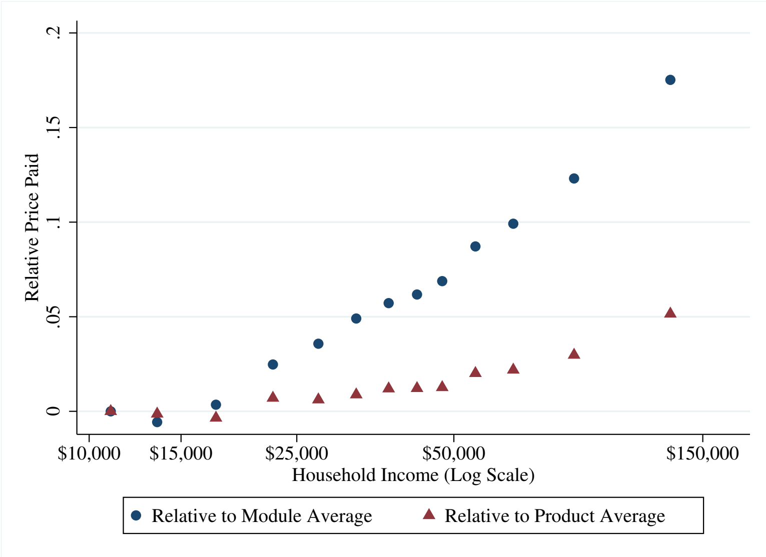
This section draws on the Nielsen HMS and RMS data described above to document two stylized facts. Taken together, these facts demonstrate the empirical patterns behind the main results of the paper. The first also serves to motivate the theoretical framework presented in Section 4 below.

3.1 High-Income Households Purchase Different, More Expensive, Products than Low-Income Households

Figure 1 shows that high-income households pay more than low-income households for the same type of products. The level of each circle shows how much more households in each Nielsen income category pay per unit for products within a module than households in the lowest income category, earning between \$10,000 and \$12,000. These relative prices are measured in a regression of log unit price paid against income category dummies and module fixed effects, controlling for other demographics with dummies for household size, marital status, race, Hispanic origin, and male and female head-of-household education and age. There is a distinct upward slope, with households in the upper-most income category paying approximately 17 percent more for products in the same module than households in the lowest income category. This could be either because high-income households are paying more for the same products within a module or because they are purchasing different, more expensive products. The following result suggests that the latter effect dominates.

The level of each triangle in Figure 1 shows how much more households in each Nielsen income category pay for the same product, relative to households in the lowest income category, measured in the same regression as described above but with product, instead of module, fixed effects. The slope of the log unit price paid controlling for product fixed effects is positive but much smaller than the slope of the log unit paid only controlling for module fixed effects. High-income households do pay more for the same products but, consistent with Broda et al. (2009), most of this gradient is explained by the fact that they are buying different products that are sold at higher prices to all consumers.

Figure 1: Average Log Price Paid by Household Income Category



Notes: This figure plots the average unit price paid by Nielsen household panelists at different income levels relative to the unit price paid by all households for either the same product or products in the same module. Relative price paid is the coefficient on a household income dummy in a regression of the log unit price paid by a household for a product in a month on module or product fixed effects and demographic controls. The relative price paid by each household income category is plotted against the mid-point of the bounds of the reported incomes for that category for all but the highest "income greater than $100,000" category, whose relative price paid is plotted at $130,000.

3.2 Stores in Wealthier Markets Offer More Products that are Purchased by High-Income than by Low-Income Households at Slightly Lower Relative Prices

Figure 2 shows that the products favored by high-income households are more likely to be available and sold at lower prices in markets with higher per capita income, relative to the products favored by low-income households. The figure is constructed using two indexes. First, a variety index V_c^k that measures the extent to which a market c offers the products favored by income group k relative to other markets. The variety index in market c for an income group k, V_c^k, is defined as the share of expenditure that HMS panelists that belong in income group k but are not in market c allocate to the products available in market c, or

$$V_c^k = \sum_{g \in \mathbf{G}_c} \left(\frac{v_{kcg}}{\sum_{g' \in \{\mathbf{G}_{c'}\}} v_{kcg}} \right)$$

where $\mathbf{G}_c$ denotes the set of products g available in market c and v_{kcg} denotes the amount that HMS panelists in income group k that are *not* in market c spend on product g in 2012.[15] Second, a simple price index P_c^k equal to the weighted average relative price charged in CBSA c, using

[15] Specifically, $v_{kcg} = \sum_{i \in \{\mathbf{I}_{kc'}\}_{c' \neq c}} v_{ig}$ where $\mathbf{I}_{kc'}$ denotes the set of HMS panelists i in size-adjusted income decile k observed in market c' and v_{ig} denotes the expenditure of HMS panelist i on product g in 2012.

income group k-specific expenditures for weights:

$$P_c^k = \sum_{g \in \mathbf{G}_c} \left(\frac{p_{cg}}{p_g} \right)^{\frac{v_{kcg}}{\sum_{g \in \mathbf{G}_c} v_{kcg}}}$$

where p_{cg} is the sales-weighted average price charged for product g in CBSA c in 2012 and p_g is the sales-weighted average price charged for product g nationally in 2012.

Figure 2a plots the gap in the variety index between the top and bottom income decile ($V_c^{10} - V_c^1$) in each CBSA against log CBSA per capita income. It reveals a statistically-significant correlation between the city wealth and product availability: the consumption opportunities in high-income cities are skewed towards those products that are consumed more heavily by high-income consumers relative to those consumed more heavily by low-income consumers. For example, around 1.2 percentage points more of the top income decile's expenditure share than that of the bottom income decile is represented in the sample for the wealthiest city, Bridgeport-Stamford-Norwalk, CT (BRI), while 1 percentage point less is represented in the sample for the poorest city, El Paso, Texas (ELP). To put these differences into context, Appendix Figure A.4 shows that wealthy cities offer greater variety of products for all income deciles, but the variety index for the top income decile increases with log per capita income at over twice the rate that the variety index for the bottom income decile increases (2.2 vs. 1.0).

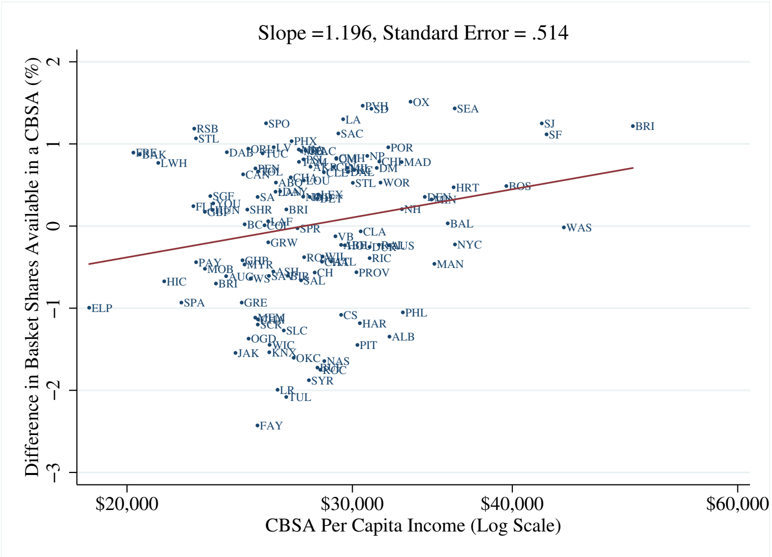
Figure 2b shows how the gap in the average relative price faced by high- and low-income households for the products they consume more of ($P_c^{10} - P_c^1$) varies across CBSAs with different per capita income. The plot shows a noisier relationship. Stores in high-income CBSAs tend to charge less for the products that high-income households purchase more of (relative to low-income households) than stores in low-income CBSAs, but this difference is small relative to the rate at which the price of bundles favored by both high- and low-income households increases with CBSA income.[16]

Table 2 replicates this analysis comparing the products available and price charged across individual grocery stores, rather than across CBSAs. Panel A compares availability patterns across stores. In column [1], we see that, in aggregate, stores in higher-income neighborhoods offer more of the products high-income households purchase more of. These availability patterns are stronger looking across stores within the same CBSA, in column [3], than across stores in CBSAs with different per capita incomes, in column [5]. In all three cases, the availability patterns are less than half as large when looking across stores in the same retail chain. The patterns in price levels, shown in Panel B, are similar, also favoring high-income consumers in higher-income neighborhoods and CBSAs, with less variation looking within chain than across

[16] Appendix Figure A.4 shows that the hedonic price indexes for both high- and low-income households increase sharply with CBSA per capita income: the semi-elasticity of the price index with respect to CBSA per capita income is 4.1 for the top decile relative to 4.9 for the bottom decile.

Figure 2: Difference in the Availability and Relative Price of High-Income and Low-Income Baskets Across CBSAs

a. Availability



b. Relative Price

Notes: Figure a. plots CBSA-level data for the difference between the expenditure shares of high-income Nielsen HMS panelists represented in the CBSA product set and the expenditure share of low-income panelists represented in that product set against CBSA per capita income. The panelist expenditure shares are calculated for 2012 and are CBSA-specific, in that they exclude the expenditures of any panelists residing in the CBSA whose availability is being measured. Figure b. plots CBSA-level data for the difference between the average price level faced by consumers in the top income decile and the average price level faced by households in the bottom income decile against CBSA per capita income. The price level in each CBSA for a given income decile is calculated as the weighted average log of the ratio between the price a product is sold for in a CBSA relative to the price that product is sold at in the national sample where weights are defined as the value of the purchases of that product made by households in the respective income decile in the Nielsen household-level panel. Panelists are defined as high- (or low-) income if their size-adjusted income falls in the top (bottom) decile of panelist incomes. The products available and prices charged in each CBSA are defined as the set of products sold and average unit prices charged in a random sample of 50 Nielsen stores in a given CBSA in 2012. The plots show the mean availablility share and price indexes calculated in 100 bootstrap iterations of this sampling procedure. CBSA income is household income adjusted for size using a square-root equivalence scale. The marker labels for each CBSA are acronyms linked to the full CBSA name in Appendix A.4.

chains. The only exception here is that the relative price charged for products that high-income consumers favor is less correlated with local income across stores in different neighborhoods of the same CBSA (column [3]) than across neighborhoods both within and across CBSAs (column [1]). Consistent with chain-level pricing, this correlation falls almost to zero when looking within chain and CBSA (column [4]). In effect, the spatial differences in product availability and prices documented in this paper can be attributed primarily to variation in store location and product distribution patterns across chains, and less to variation in product distribution patterns across stores within the same chain.

This section has established that there are large systematic differences in product availability between wealthy and poor markets and that these differences are correlated with the purchase behavior of high- and low-income households. Stores in wealthy markets also charge relatively less for products that the top income decile's consumption basket than the bottom income decile's consumption basket, but these differences are small relative to the rate at which prices increase with market income for both income deciles. Whether the variety benefits of wealthy markets outweigh the higher prices charged in these markets to make the variety-adjusted price index higher or lower for any given income group is an empirical question that cannot be answered with the ad hoc variety and price indexes studied above. The structural analysis below will quantify how much high- and low-income households gain from the relative abundance of these products available in wealthy cities and neighborhoods across the U.S. and the extent to which these variety gains offset the higher prices charged in these locations for households in each income group.[17]

4 Model

This section introduces the demand system I use to study why high-income households purchase different products to low-income households and at different prices. This framework also forms the basis of the price indexes that summarize how high- and low-income households value the prices and products available to them in different markets.

4.1 Notation

Figure 3 shows how consumers choose to allocate expenditures. At the upper-most level, a consumer i spends W on a set of grocery products, denoted $\mathbf{G}$, and Z on a set of other goods,

[17] Handbury and Weinstein (2014) find that the variety benefits of larger cities, which also tend to be wealthier, outweigh the additional costs of the higher prices observed in these locations. In both papers, the benefits of having a greater number of products available in a market depends on the estimated elasticity of substitution between products. Here, the benefits of having a mix of products biased towards one's (non-homothetic) tastes will further depend on the estimated strength of that non-homotheticity in demand, modeled in Section 4 below.

Table 2: Difference in the Availability and Relative Price of High-Income and Low-Income Baskets Across Stores

Panel A: Availability

	Dependent Variable: Difference in Basket Shares (%)					
	[1]	[2]	[3]	[4]	[5]	[6]
Ln(Local Per Capita Income)	2.12***	0.70***	2.47***	1.07***		
	(0.39)	(0.15)	(0.24)	(0.091)		
Ln(CBSA Per Capita Income)					1.87***	0.49***
					(0.44)	(0.18)
CBSA Fixed Effects	No	No	Yes	Yes	No	No
Chain Fixed Effects	No	Yes	No	Yes	No	Yes
Number of CBSAs	-	-	-	-	691	691
Observations	9,019	9,019	8,849	8,849	9,019	9,019
adj. R^2	0.15	0.79	0.56	0.89	0.08	0.78

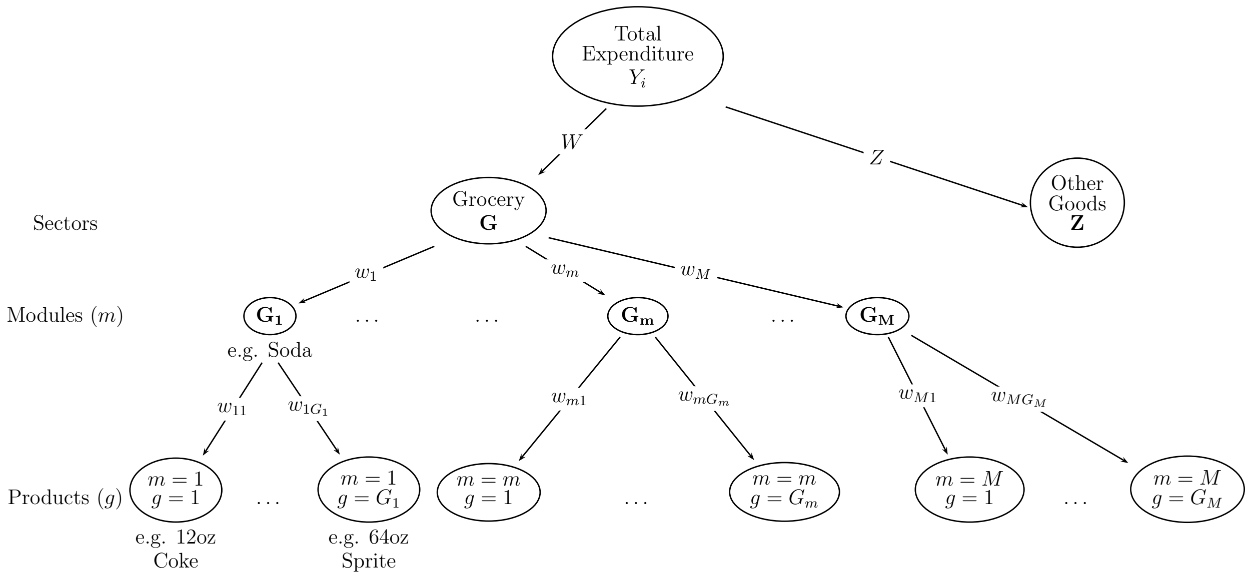
Panel B: Relative Price

	Dependent Variable: Difference in Price Level (%)					
	[1]	[2]	[3]	[4]	[5]	[6]
Ln(Local Per Capita Income)	-1.30***	-0.36***	-0.58***	-0.067		
	(0.18)	(0.084)	(0.15)	(0.090)		
Ln(CBSA Per Capita Income)					-1.44***	-0.46***
					(0.23)	(0.12)
CBSA Fixed Effects	No	No	Yes	Yes	No	No
Chain Fixed Effects	No	Yes	No	Yes	No	Yes
Number of CBSAs	-	-	-	-	691	691
Observations	9,019	9,019	8,849	8,849	9,019	9,019
adj. R^2	0.18	0.72	0.51	0.79	0.14	0.72

Notes: *** p< 0.01, ** p<0.05, * p<0.10; standard errors, clustered by CBSA, are in parentheses. The table reports the results of fixed-effect regressions. In the Panel A, the dependent variable is the difference between the share of the high-income Nielsen HMS panelist expenditures represented in the set of products sold by a store in 2012 and the share of low-income panelist expenditures represented in that same product set. In Panel B, the dependent variable is the difference between the average price level faced by consumers in the top income decile and the average price level faced by households in the bottom income decile against local per capita income. The price level in each store for a given income decile is calculated as the weighted average ratio between the price a product is sold for in a store relative to the price that product is sold at in the national sample where weights are defined as the value of the purchases of that product made by households in the respective income decile in the Nielsen household-level panel. In each column, this dependent variable is regressed against the log per capita income of the neighborhood (in columns 1 through 4) or CBSA (in columns 5 and 6) where the store is located, as well as chain fixed effects in columns 2, 4, and 6. The number of observations decreases when introducing CBSA fixed effects because not all stores are located in CBSAs.

denoted $\mathbf{Z}$, subject to the budget constraint $W + Z \leq Y_i$. I do not explicitly model this upper-level expenditure allocation decision, but it is crucial in one respect: preferences over grocery products are non-homothetic because they depend on aggregate non-grocery expenditures. This is generically the case if optimal non-grocery expenditures are normal.[18]

Figure 3: Consumer Choices



This paper focuses on the choices that consumers make within the grocery sector; that is, how consumers allocate their grocery expenditure W between product modules, $\mathbf{M} = \{1, ..., M\}$, and their module expenditure w_m between the varieties of grocery products in module m, $\mathbf{G_m} = \{1, ..., G_m\}$, for each module m. A consumer chooses to spend some w_{mg} on each product g in module m, purchasing $q_{mg} = w_{mg}/p_{mg}$ units of the product at a unit price p_{mg}. I denote the set of observed grocery prices and purchase quantities for module m as $\mathbb{P}_m = \{p_{mg}\}_{g \in \mathbf{G_m}}$ and $\mathbb{Q}_m = \{q_{mg}\}_{g \in \mathbf{G_m}}$, respectively. $\mathbb{P}$ and $\mathbb{Q}$ are the unions of these price and quantity sets over all modules. A consumer's across-module and within-module expenditure allocation decisions are linked by the fact that they cannot allocate more than their total module expenditure, w_m, between products $g \in \mathbf{G_m}$; that is, $\sum\limits_{g \in \mathbf{G_m}} w_{mg} = w_m$.

4.2 Consumption Utility

I model consumer demand for the products in $\mathbf{G}$ using a combination of Cobb-Douglas and log-logit preferences. A consumer i's utility from grocery consumption, conditional on their

[18]Formally, preferences cannot depend on expenditures, so Z is rather an aggregate of non-grocery consumption. In Appendix A, I solve for an implicit restriction on utility and prices under which the optimal non-grocery expenditure, Z_i^*, will be increasing in income. I cannot show that this restriction holds generally, but am instead able to show that it holds in the data.

non-grocery expenditure Z, is a Cobb-Douglas aggregate over consumer-specific module-level utilities:

$$U_{iG}(\mathbb{Q}, Z) = \prod_{m \in M} \left(u_{im}\left(\mathbb{Q}_m, Z\right)\right)^{\lambda_m} \tag{1}$$

where $\lambda_m \in (0, 1)$ are module-level expenditure weights and $\sum_{m \in M} \lambda_m = 1$.

Consumer i's utility from consumption in module m, conditional on their non-grocery expenditure Z, is equal to the sum of their consumer-specific product-level utilities:

$$u_{im}\left(\mathbb{Q}_m, Z\right) = \sum_{g \in \mathbf{G_m}} u_{img}(\mathbb{Q}_m, Z) \tag{2}$$

where consumer i's utility from consuming q_{mg} of product g in module m, conditional on their non-grocery expenditure Z, is defined as:

$$u_{img}(Z) = q_{mg} \exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img}) \tag{3}$$

where β_{mg} is the quality of product g in module m and ε_{img} is the idiosyncratic utility of consumer i from product g in module m drawn from a type I extreme value distribution. $\gamma_m(Z)$ and $\mu_m(Z) > 0$ are weights that govern the extent to which consumers with non-grocery expenditure Z care about product quality and their idiosyncratic utility draws.[19]

4.2.1 Functional Forms

Before proceeding, it is worth making three observations about the general functional forms assumed above. First, the Cobb-Douglas utility function governing the cross-module substitution patterns implies that consumers will optimally consume a positive amount in each module. In the data for 2012, the typical household buys products in around one third of sample modules. This purchase behavior could reflect that households are, on average, consuming small quantities of products in some modules and, therefore, purchase the product so infrequently that we

[19]The log-logit utility function defined in equations (2) and (3) is a generalization of a utility function used by Auer (2010) to theoretically derive the effects of consumer heterogeneity on trade patterns and the welfare gains from trade.

do not observe a purchase over the time period that they are in the sample.[20,21]

Second, the assumption that module utility is additive in product utilities that themselves are proportional to random draws from a continuous (type I extreme value) distribution implies that households allocate all of their module expenditure to a single product (the product that maximizes their marginal utility from expenditure, $\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})/p_{mg}$). This matches the discrete-continuous behavior observed in the data: conditional on purchasing any products in a module in a month, households typically only purchase one product.

Finally, the log-logit function governing preferences within modules yields the same Marshallian demand function for a set of consumers as the nested-CES utility function for a representative consumer with non-grocery expenditure Z and an elasticity of substitution between products equal to one plus the inverse of the idiosyncratic utility draw weight, i.e., $\sigma_m(Z) = 1 + 1/\mu(Z)$. This link provides a natural analytic approximation for the relative utility that consumers with the discrete-continuous preferences described above face across markets offering different choice sets. The log-logit functional form also implies that, conditional on non-grocery expenditure, preferences are weakly-separable between modules. I exploit these features in the empirical strategy presented in Section 5.1 below.

4.2.2 Non-Homotheticities

Consumers get utility from consuming quantity q_{mg} of a product g, scaled up by exponents of product quality, β_{mg}, and idiosyncratic utility, ε_{img}. Preferences will be non-homothetic when at least one of the weights on these scalars, $\gamma_m(Z)$ or $\mu_m(Z)$, varies with non-grocery expenditure and, as discussed above, this expenditure varies with income. In order to interpret how these weights vary with income empirically, I make further functional form assumptions.

I interpret $\gamma_m(Z)$ to be the valuation for product quality, β_{mg}, for product g in module m shared by consumers with non-grocery expenditure Z. I assume that $\gamma_m(Z)$ is log-linear in Z

[20] In this scenario, households make purchases in all modules in expectation. The moments used to estimate the model parameters are based on individual household product selections within modules, conditional on their making a purchase in a given module, and expected store sales, i.e., the purchases of many households that shop in a store. The fact that some households do not purchase products in certain modules during a given period will be reflected in the fact that these modules have low within-store sales shares, and explained by the fact that the products in these modules are, on average, either more expensive or lower quality, relative to products in other modules. Models that reflect these more realistic cross-module consumption patterns, either by accounting for dynamic purchase behavior (see, *e.g.*, Hendel (1999); Dube (2004)) or explicitly modeling consumer's discrete-continuous preferences over modules (see, *e.g.*, Song and Chintagunta (2007); Pinjari and Bhat (2010)), would be difficult to estimate given the dimensions of the problem that this paper addresses.

[21] The Cobb-Douglas utility function is also restrictive in other respects. Supplemental Appendix B presents the model, estimation procedure, and results under the more flexible assumption of CES utility across modules. The results are similar to the baseline Cobb-Douglas model assumed here because the estimated cross-module substitution elasticities are close to one.

with a module specific slope, γ_m, such that:

$$\gamma_m(Z) = 1 + \gamma_m \ln(Z) \tag{4}$$

A consumer's valuation for product quality in module m is increasing in Z when $\gamma_m > 0$.

I employ a revealed preference approach to estimate the product quality β_{mg} parameters as the average willingness to pay for product g in module m across all consumers. The idea here is that product g in module m is estimated as having high quality, β_{mg}, relative to that of another product $\tilde{g}$ in the same module m, $\beta_{m\tilde{g}}$, when a set of consumers facing the same price for both products spends a higher share of their expenditure on product g than on product $\tilde{g}$. All consumers agree on this distribution of product qualities but, for $\gamma_m > 0$, consumers who spend more on non-grocery items place a greater weight on product quality, relative to quantity, in selecting which product to purchase in a module. Since Z is normal, a positive γ_m implies that high-income consumers spend a disproportionate amount of their module expenditures on higher quality products, relative to low-income consumers.

This form of non-homotheticity is common in the international trade literature where, for example, Fajgelbaum et al. (2011) show the theoretical implications of non-homothetic demand with a model that allows for complementarities between product quality and expenditure on a non-differentiated outside good. These complementarities imply that the elasticity of demand for quality is increasing with income, as in Hallak (2006) and Feenstra and Romalis (2014), who calculate cross-country price indexes similar to those estimated below.

The within-module utility function defined in equations (2) and (3) is also non-homothetic through the weight, $\mu_m(Z)$, on the idiosyncratic utility, ε_{img}. These idiosyncratic utility weights govern the dis-utility from consuming products that are horizontally differentiated from the consumer's ideal type of product, or the extent to which consumers find the available products substitutable with their ideal. I assume that the inverse of the idiosyncratic utility draw weight for module m is log linear in non-grocery expenditures:

$$\frac{1}{\mu_m(Z)} = \sigma_m(Z) - 1 \equiv \alpha_m^0 + \alpha_m^1 \ln(Z) \tag{5}$$

where recall that $\sigma_m(Z)$ reflects the elasticity of substitution between products in module m for a representative consumer with non-grocery expenditure Z. For $\alpha_m^1 < 0$, $\sigma_m(Z)$ decreases with Z such that consumers with high non-grocery expenditures find the available products less substitutable with each other and their ideal product and will, therefore, have a higher willingness to pay for the product closest to their ideal than consumers with low non-grocery expenditures. That is, for Z normal, $\alpha_m^1 < 0$ implies that consumers' elasticity of substitution between products within a module and their tendency to switch between products in response to relative price changes is decreasing in consumer income.

This form of non-homothetic price sensitivity is also similar to those used in recent international trade models. Hummels and Lugovskyy (2009), for example, develop a Lancaster ideal variety utility function where the dis-utility from distance between a product and a consumer's ideal type is an increasing function of their consumption quantity q^{γ} for $\gamma \in [0,1]$. This weight implies an income-specific price elasticity in a similar manner to the idiosyncratic utility weights, $\mu_m(Z)$, above.[22]

4.3 Individual Utility Maximization Problem

The grocery utility function defined in equations (1)-(3) is specific to the individual through a consumer's income, their non-grocery expenditure, and their idiosyncratic utility draws. I assume that consumers draw an idiosyncratic utility ε_{img} for each product $g \in \mathbf{G}$ prior to making their purchase decision. Consumers then solve for their optimal grocery consumption bundle for a given non-grocery expenditure level Z by maximizing grocery utility subject to budget and non-negativity constraints:

$$(6) \qquad \sum_{m \in \mathbf{M}} \sum_{g \in \mathbf{G_m}} p_{mg} q_{mg} \leq Y_i - Z \quad \text{and} \quad q_{mg} \geq 0 \, \forall mg \in \mathbf{G}$$

The solution to this problem is a vector of optimal product selections (one for each module), $\mathbf{g_i^*}(Z) = (g_{i1}^*(Z), ..., g_{iM}^*(Z))$ and module-level expenditures, $\mathbf{w_i^*}(Z) = (w_{i1}^*(Z), ..., w_{iM}^*(Z))$. The optimal product selections (derived in Appendix C.1) are

$$(7) \qquad g_{im}^*(Z) = \arg\max_{g \in \mathbf{G_m}} \left(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img}\right)/p_{mg}$$

and, given the Cobb-Douglas assumption, the module-level expenditures are

$$(8) \qquad w_{im}^*(Z) \quad = \quad (Y_i - Z)\lambda_m$$

Plugging these optimal product choices and module expenditures into the direct utility function

[22] Macro-economists have found alternative models to be empirically relevant for explaining differences in the prices paid by high- and low-income households. These models appear to be less relevant in the Nielsen data, so it is unlikely that ignoring them biases the aggregate estimates found below. The cross-income differences in search costs and shopping behavior explored in Simonovska (2015) could, in theory, enable low-income households to mitigate the high prices in wealthy cities at a lower cost than high-income households. However, Figure 1 shows that the cross-income differences in prices paid for identical items purchased in different stores or at different sale/non-sale periods are relatively small compared to the unit expenditure differences attributable to the fact that high- and low-income consumers are buying entirely different products. I also find no evidence that high-income consumers purchase more varieties of bar-coded products than low-income consumers, as would be the case in a hierarchic demand model like that use to explain Indian household consumption in Li (2021) or the translated additive-log utility function used in Simonovska (2015).

defined in equations (1)-(3), I obtain the indirect utility of consumer i from grocery consumption in a market offering prices and products summarized in the vector $\mathbb{P}$:

$$V(\mathbb{P}, Y_i, Z, \varepsilon_{\mathbf{i}}) = \frac{(Y_i - Z)}{P(\mathbb{P}, Z, \varepsilon_{\mathbf{i}})} \tag{9}$$

where $P(\mathbb{P}, Z, \varepsilon_i)$ is a Cobb-Douglas price index over the grocery products that a consumer i optimally consumes in each module:

$$P(\mathbb{P}, Z, \varepsilon_{\mathbf{i}}) = \prod_{m \in \mathbf{M}} \left(\max_{g \in \mathbf{G_m}} \left(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img}\right) / p_{mg} \right)^{\lambda_m} \tag{10}$$

5 Empirical Strategy

A key goal of this paper is to characterize how consumers at different income levels value the different products and prices available to them across different markets in the U.S.. In this section, I first derive the income- and city-specific price indexes I use to measure this variation. These indexes require two key components: vectors of the prices that provide comparable representations of the prices and product variety available in different U.S. cities, and estimates for model parameters that govern consumer's perceptions of these price vectors. In the remainder of the section describes how I use the Nielsen data to obtain each of these components.

5.1 Measuring Relative Utility Across Markets

Section 4.3 above solved for the indirect utility of a consumer from grocery consumption in a generic market offering a vector of prices $\mathbb{P}$. To compare the utility consumers get from the prices and products available to them in different markets, I now introduce a market subscript to equation (9), writing the indirect utility of a consumer i in market t as

$$V(\mathbb{P}_t, Y_i, Z_{it}, \varepsilon_{\mathbf{i}}) = \frac{(Y_i - Z_{it})}{P(\mathbb{P}_t, Z_{it}, \varepsilon_{\mathbf{i}})} \tag{11}$$

where the set of prices and products available to household i, $\mathbb{P}_t = \{p_{mgt}\}_{g \in \mathbf{G}_t}$, and their optimal non-grocery expenditures, Z_{it}, are both allowed to vary across markets.

This indirect utility function is consumer-specific in three ways: it depends on a consumer's income, Y_i, on their optimal non-grocery expenditures, Z_{it}, and on their idiosyncratic utility draws, $\varepsilon_{\mathbf{i}}$. To study the systematic variation in utility across consumers earning different incomes, I abstract from any variation in non-grocery expenditures Z_{it} and/or idiosyncratic utility draws $\varepsilon_{\mathbf{i}}$ that is uncorrelated with income. The idiosyncratic utility $\varepsilon_{\mathbf{i}}$ draws are, by definition, uncorrelated with consumer income Y_i. The most direct way to abstract from this random variation would be to take the expectation of the indirect utility defined in equation (11) over the idiosyncratic draws. Unfortunately, there is no analytic solution to this problem, and numerical solutions are computationally intensive. Instead, I approximate the relative utility of households at a given income level across different markets with the relative utility of an income-specific representative consumer at the same income across the same markets.

The representative consumer's utility from consuming a grocery bundle $\mathbb{Q}$ is a weighted geometric mean of module-level CES utilities conditional on their non-grocery expenditure Z defined as:

$$U_G^{CES}(\mathbb{Q}, Z) = \prod_{m \in M} \left[\sum_{g \in \mathbf{G_m}} \left[q_{mg} \exp(\beta_{mg} \gamma_m(Z)) \right]^{\frac{\sigma_m(Z)-1}{\sigma_m(Z)}} \right]^{\left(\frac{\sigma_m(Z)}{\sigma_m(Z)-1}\right)\lambda_m}, \tag{12}$$

where q_{mg}, β_{mg}, $\gamma_m(Z)$, $\sigma_m(Z)$, and λ_m take the same definitions as in the nested log-logit utility function presented in Section 4 above.[23] The indirect utility of this representative consumer from income Y_i and prices and products $\mathbb{P}_t$, $V^{CES}(\mathbb{P}_t, Y_i)$, takes a similar form to the indirect utility of the idiosyncratic consumer provided in equation (11) above. It can also be expressed as the ratio of the consumer's grocery expenditure to a price index that summarizes the consumer's marginal utility from expenditure given the prices and products available in the market:

$$V^{CES}(\mathbb{P}_t, Y_i, Z_{it}) = \frac{(Y_i - Z_{it})}{P^{CES}(\mathbb{P}_t, Z_{it})}, \tag{13}$$

where

$$P^{CES}(\mathbb{P}_t, Z_{it}) = \prod_{m \in \mathbf{M}} \left(\left[\sum_{g \in \mathbf{G_{mt}}} \left(\frac{p_{mgt}}{\exp(\beta_{mg} \gamma_m(Z_{it}))} \right)^{(1-\sigma_m(Z_{it}))} \right]^{\frac{\lambda_m}{1-\sigma_m(Z_{it})}} \right)$$

for p_{mgt} equal to the unit price at which product g in module m is sold in market t.

To summarize this indirect utility function across households so that it varies with i only through income, Y_i, I approximate household non-grocery expenditures by assuming that non-grocery expenditures, Z_{it}, vary only with household income, Y_i, such that $Z_{it} = Z(Y_i)$.[24] Under

[23]In Supplemental Appendix B, I show that this income-specific, Cobb Douglas-nested CES utility function yields identical within-grocery budget shares as the Cobb Douglas-nested log-logit utility function that I estimate.

[24]Theoretically, this assumption could be violated since consumers at each income level may optimally choose different aggregate expenditure allocations across cities to suit the different grocery and non-grocery prices they face in these locations. Empirically, however, I observe that the relationship between non-grocery expenditures and income is surprisingly consistent across cities. Appendix Figure B.1 demonstrates that households earning higher incomes spend a smaller share of their income on grocery products. Within income groups, however, the average grocery expenditure share does not vary much across cities and, in particular, it does not vary systematically with city income.

this assumption, we can express the consumer's indirect utility as a function of market prices, $\mathbb{P}_t$, and consumer income, Y_i alone:

$$V^{CES}(\mathbb{P}_t, Y_i) = \frac{(Y_i - Z(Y_i))}{P^{CES}(\mathbb{P}_t, Z(Y_i))}, \tag{14}$$

where

$$P^{CES}(\mathbb{P}_t, Z(Y_i)) = \prod_{m \in \mathbf{M}} \left(\left[\sum_{g \in \mathbf{G_{mt}}} \left(\frac{p_{mgt}}{\exp(\beta_{mg}\gamma_m(Z(Y_i)))} \right)^{(1-\sigma_m(Z(Y_i)))} \right]^{\frac{\lambda_m}{1-\sigma_m(Z(Y_i))}} \right) \tag{15}$$

In particular, a consumer's relative indirect utility across two markets t and t' is equal to the inverse of the relative price indexes they face across the same markets:

$$\frac{V(\mathbb{P}_t, Y_i)}{V(\mathbb{P}_{t'}, Y_i)} = \frac{P^{CES}(\mathbb{P}_{t'}, Z(Y_i))}{P^{CES}(\mathbb{P}_t, Z(Y_i))} \tag{16}$$

That is, the magnitude of the price index a consumer with income Y_i faces in market t relative to the price index they face in market t' indicates how much lower (or higher) the consumer's grocery utility is in market t relative to market t'. The remainder of this section outlines how I obtain the two key inputs for these price indexes: market-specific price vectors and demand parameters.[25]

5.2 Inferring Prices and Product Availability

The first input to the price index defined in equation (15) is a market-specific price vector, $\mathbb{P}_t$, representing the set of prices and products available to consumers in a market t. I calculate price indexes comparing grocery costs across two types of markets in 2012: CBSAs and stores.

[25]Note that this approach to measuring income-specific spatial price indexes is different from the approach that Broda and Romalis (2009) developed to calculate income-specific inflation with the same Nielsen household-level data. Broda and Romalis (2009), and subsequent papers by Argente and Lee (2016) and Jaravel (2018), use the Feenstra (1994) methodology to calculate price indexes that are exact to a nested-CES utility function similar to the one above, but with two key differences. The Broda and Romalis (2009) approach is more restrictive in that the authors do not allow the substitution elasticities, σ_m in the framework above, to vary with income. It is, however, more flexible in implicitly allowing for households at different income levels to have entirely different revealed preferences (β_{mg}s) for products. In the model presented here, households agree on the qualities of products and only the willingness to pay for quality varies with household income. The additional structure imposed on the relationship between perceived quality and income in this paper, as well as in more recent work by Feenstra and Romalis (2014), provides a clearer economic interpretation for the cross-income differences in the relative costs measured here relative to those measured in Broda and Romalis (2009). The Feenstra and Romalis (2014) approach is similar to mine in that the authors estimate the parameters of the underlying utility function and use these estimates to adjust prices for product quality. While the resulting price indexes are not income-specific, they are based on a utility function that is non-homothetic in demand for quality in the same way as the utility function presented above.

I proxy for the set of prices and products available to consumers in each CBSA in 2012 using the set of products and unit prices represented in the 2012 sales of a random sample of the RMS participating retailers located in that CBSA, as described in Section 2 above. I proxy for the prices and products available to consumers in individual grocery stores in 2012 using the set of products and unit prices observed in the sales of each establishment in 2012.

5.3 Parameter Estimation

The second set of inputs into the price index defined in equation (15) are model parameters that characterize how consumers value the products and prices available to them in a market, and how this valuation varies with consumer income. I denote this set of parameters using a vector θ defined as

$$\theta = \{(\theta_1, \ldots, \theta_M)\}$$

where $\theta_m = \{\alpha_m^0, \alpha_m^1, \beta_{m1}, \ldots, \beta_{mG_m}, \gamma_m, \lambda_m\}$. I estimate these parameters in two stages. The first stage identifies the parameters that govern the relative shares households spend on different products within each module; that is, all components of θ_m except for the quality parameter $\beta_{m\bar{g}_m}$ of a module-specific base product $\bar{g}_m$ and the Cobb-Douglas module weight, λ_m. I denote this set of parameters by $\theta_1 = \{\theta_{1m}\}_{m \in \mathbf{M}}$ where

$$\theta_{1m} = \left\{\alpha_m^0, \alpha_m^1, \gamma_m, \left\{\tilde{\beta}_{mg}\right\}_{g \in \mathbf{G}_m}\right\}$$

for each module $m \in \mathbf{M}$ and tildes denote that a variable has been differenced from the respective value for the outside product in each module, $\bar{g}_m$ (e.g., $\tilde{\beta}_{mg} = \tilde{\beta}_{mg} - \tilde{\beta}_{m\bar{g}_m}$). The estimation routine follows Berry et al. (2004) and is described further below. In the second stage, I fit the Cobb-Douglas module weights, λ_m, to the sales share of each module m in the store-level data.

Under the assumption of Cobb-Douglas demand over modules, the remaining parameters – the base-product qualities, $\{\beta_{m\bar{g}_m}\}_{m=1,\ldots,M}$ – are not identified. Without these base quality parameters, I cannot measure how grocery costs vary across households with different incomes in the same city. I can, however, measure how grocery costs vary across cities within each income group and, therefore, importantly can ascertain how grocery costs vary across cities differently for households at different income levels.[26]

[26]To see this, notice that we can re-write the price index faced by a representative household with income Y_i in market t defined in (15) above as a market-invariant aggregate of base product qualities, $\mathbb{B}(Z(Y_i))$, and a variant of the price index in equation (15) calculated using normalized product quality $\tilde{\beta}_{mg}$ in place of absolute product quality β_{mg}; that is,

$$P^{CES}(\mathbb{P}_t, Z(Y_i)) \quad = \quad \mathbb{B}(Z(Y_i))\tilde{P}^{CES}(\mathbb{P}_t, Z(Y_i)) \qquad \text{where} \qquad \mathbb{B}(\text{Z}(Y_i)) = \prod_{m \in \mathbf{M}} (\exp(\beta_{m\bar{g}_m}\gamma_m(Z(Y_i))))^{\lambda_m}.$$

5.3.1 Within-Module Estimation Methodology

To estimate the parameters that govern the within-module substitution patterns, I employ a GMM procedure to fit two sets of predicted moments to their data analogs. These moments are (1) store-level product sales shares and (2) the covariance of the prices and estimated qualities of the products purchased by each household with household income. The moment conditions and variation that identifies each parameter is described further below.

Estimation Procedure Given the distributional assumption on ε_{img}, the conditional probability of purchasing product g in module m for a household with non-grocery expenditure Z_i and facing a vector of prices $\mathbb{P}$ takes the familiar multinomial logit form:

$$P_{mg}(Z_i, \mathbb{P}, \theta_m) = \frac{\exp\left[\alpha_{im}\left(\gamma_{im}\beta_{mg} - \ln p_{mg}\right)\right]}{\sum_{g' \in \mathbf{G_m}} \left(\exp\left[\alpha_{im}\left(\gamma_{im}\beta_{mg'} - \ln p_{mg'}\right)\right]\right)} \tag{17}$$

where $\alpha_{im} = (\alpha_m^0 + \alpha_m^1 \ln Z_i)$ and $\gamma_{im} = (1 + \gamma_m \ln Z_i)$.

The first set of moments fits predicted product market shares to those observed in the RMS data. I calculate these sales shares using data aggregated to the CBSA-month level to mitigate biases associated with low and zero sales shares. Accordingly, I adjust the standard purchase probability expressed in equation (17) to reflect time-varying CBSA-specific pricing and promotion activity:

$$P_{mg}(Z_i, \mathbb{P}_{st}, \theta_m, \xi_t) = \frac{\exp\left[\alpha_{im}\left(\gamma_{im}\beta_{mgt} - \ln p_{mgt}\right)\right]}{\sum_{g' \in \mathbf{G_{mt}}} \left(\exp\left[\alpha_{im}\left(\gamma_{im}\beta_{mg't} - \ln p_{mg't}\right)\right]\right)}$$

where $\beta_{mgt} = \beta_{mg} + \xi_{mgt}$ and ξ_{mgt} is a transitory taste shock for product g in CBSA-month market t, demeaned from the fixed product quality parameter, β_{mg}. The fixed product quality parameter refers to characteristics of the product that are common across CBSAs and over time, such as physical characteristics of the product itself and national recognition of the product's brand. The transitory taste shock is associated with local brand tastes and non-price promotions. In this stage of estimation, the product quality and the transitory taste shock will be identified for all but one product in each module, so will be estimated relative to the taste shock for the outside product (the set of products with average positive sales shares below the 60th percentile for all products).

The predicted sales of product g in module m in market t is then the aggregate of individual choice probabilities over the units purchased by customers at each non-grocery expenditure level:

$$
Q_{mgt}(\theta_m;\mathbb{P}_t) = \int \frac{\exp\left[\alpha_{im}\left(\gamma_{im}\beta_{mgt} - \ln p_{mgt}\right)\right]}{\sum\limits_{g'\in\mathbf{G_{mt}}} \exp\left[\alpha_{im}\left(\gamma_{im}\beta_{mg't} - \ln p_{mg't}\right)\right]} dF(Z_i|t) \tag{18}
$$

where $F(Z_i|t)$ is the distribution of non-grocery expenditures over all customers i in market t weighted by the number of module-m units each purchases.

The first set of moment conditions is constructed using the product of the transitory component of unobserved product quality, $\xi_{mgt}(\mathbf{X}_m;\theta_{1m})$, with a vector of pre-determined variables, $\mathbf{W}_{mgt}$, including product fixed effects and instruments described below:

$$
\bar{g}^1(\theta_m) = \frac{1}{n_m}\sum_{mg,t} g^1_{mgt}(\theta_m) = \frac{1}{n_m}\sum_{mg,t} \tilde{\xi}_{mgt}(\mathbf{X}_m;\theta_{1m})\tilde{\mathbf{W}}_{mgt}
$$

where n_m is the number of (product-CBSA-month) observations.

The second and third set of moment conditions respectively compare the covariance between the relative quality and unit value of the products purchased by households and their non-grocery expenditure to that predicted by the model. Following Berry et al. (2004), I fit the model's predictions for the uncentered covariance of quality and price with household non-grocery expenditure, i.e., $E(x_{mg}Z)$ for $x_{mgt} \in \left\{\tilde{\beta}_{mg}, \tilde{p}_{mgt}\right\}$, to that observed in the HMS data.

The quality-covariance moments are obtained from the difference between the average non-grocery expenditure of Nielsen panelists who purchase each product g in market t and the average non-grocery expenditure predicted by the model for households that purchase product g in market t. If $y = mg$ denotes that a household purchases a unit of product g in module m, i_{mg} denote one of the N_{mg} units purchased by sample households, and $N_m = \sum_{g\in\mathbf{G}_m} N_{mg}$, the quality-co variance moments are:

$$
\bar{g}^2(\theta_m) \approx \frac{1}{N_m}\sum_{mg} N_{mg}\beta_{mg}\left\{\frac{1}{N_{mg}}\sum_{i_{mg}=1}^{n_{mg}} Z_{i_{mg}} - E[Z|y=mg,\theta_m]\right\}
$$

I calculate $E[Z|y = mg, \theta]$ by first transforming it into an expression that depends on the model's predicted choice probabilities for each unit purchased:

$$
E[Z|y=mg,\theta_m] = \frac{\int\int ZP(y=mg|Z,\theta_m,y=mt)F(Z|m,t)G(t|y=m)}{\int Pr(y=mg,|\theta_m,y=m)G(t|y=m)}
$$

where $F(Z|m,t)$ is now the distribution of non-grocery expenditures of the households observed to be purchasing units of module-m products in market t, weighted by units purchased, and $G(t|y=m)$ is the distribution of these purchases across markets. In practice, I calculate

$$E[Z|y=mg,\theta_m] \;=\; \frac{\frac{1}{N_m}\sum_i Z_i P_{mg}(Z_i,\mathbb{P}_t,\theta_m,\xi_t)}{\frac{1}{N_m}\sum_i P_{mg}(Z_i,\mathbb{P}_t,\theta_m,\xi_t)}$$

where $N_m = \sum_{mg} N_{mg}$ is the total number of units sold and i indexes each unit purchased by a household i with non-grocery expenditure Z_i. This assumes that households receive an independent taste shock for each unit they purchase. $P_{mg}(Z_i,\mathbb{P}_t,\theta_m,\xi_t)$ is defined above in equation (17).

The price-covariance moments compare the covariance between the relative unit price paid by households for their selection and their non-grocery expenditure to that predicted by the model:

$$\bar{g}^3(\theta_m) \;\approx\; \frac{1}{N_m}\sum_i \left(Z_i-\bar{Z}\right)\sum_{s,t}\left(\tilde{p}_{imt}-E[\tilde{p}_{imt}|\theta_m]-\frac{1}{N_m}\sum_{i,t}\left(\tilde{p}_{imt}-E[\tilde{p}_{imt}|\theta_m]\right)\right)$$

where $\bar{Z} = \frac{1}{N_m}\sum_i \bar{Z}_i$ is the unit-weighted mean non-grocery expenditure of sample households. The relative unit price paid by a household i in module m in market t is defined as the difference between the unit price charge by the store for product household i selected from the weighted average unit price charged by stores in that market for products in that module: $\tilde{p}_{imt} = (p_{imgt}-\bar{p}_{mt})$, where $\bar{p}_{mt} = \sum_{g\in\mathbf{G}_{mt}} w_{mgt}p_{mgt}$ and $w_{mgt} = s_{mg}/\sum_{g\in\mathbf{G}_{mt}} s_{mg}$ is the product sales weight taken from the CBSA-level data. I calculate the predicted relative unit price paid by household i in module m in market t, as

$$E[\tilde{p}_{imt}|\theta_m] \;=\; \sum_{g\in\mathbf{G}_{mt}} \tilde{p}_{mgt}P_{mg}(Z_i,\mathbb{P}_t,\theta_m,\xi_t)$$

Estimation Procedure The three moment conditions defined above identify all of the module-specific parameters, θ_m, except for the quality parameter $\beta_{m\bar{g}_m}$ of the outside product $\bar{g}_m$ in each module. I denote this set of parameters by $\theta_1 = \{\theta_{1m}\}_{m\in\mathbf{M}}$ where

$$\theta_{1m} = \left\{\alpha^0_m,\alpha^1_m,\gamma_m,\left\{\tilde{\beta}_{mg}\right\}_{g\in\mathbf{G}_m}\right\}$$

for each module $m\in\mathbf{M}$.

The θ_1 parameters are estimated in separate non-linear GMM procedures that minimize a quadratic function over the moment conditions $\{\bar{g}^1(\theta_m),\bar{g}^2(\theta_m),\bar{g}^3(\theta_m)\}$ for each module m. I use the nested fixed-point algorithm proposed by Berry et al. (1995) to obtain the relative product quality parameters, $\left\{\tilde{\beta}_{mg}\right\}_{g\in\mathbf{G}_m}$, as a function of the three non-linear parameters for each module, $\theta^{NL}_{1m} = \{\alpha^0_m,\alpha^1_m,\gamma_m\}$. Given a guess of θ^{NL}_{1m}, I first invert the share equation for the relative product quality shocks, $\tilde{\beta}_{mgst}(\theta_{1m}^{NL}) = \beta_{mgst}(\theta_{1m}^{NL}) - \beta_{m\bar{g}_m st}(\theta_{1m}^{NL})$, that solve a system of non-linear equations equating predicted and observed demand at each market. I project $\tilde{\beta}_{mgt}(\theta_{1m}^{NL})$ on product dummies to obtain estimates for relative product quality $\tilde{\beta}_{mg}(\theta_{1m}^{NL})$. The residuals provide estimates for the transitory shocks, $\tilde{\xi}_{mgt}(\theta_{1m}^{NL}) = \tilde{\beta}_{mgt}(\theta_{1m}^{NL}) - \tilde{\beta}_{mg}(\theta_{1m}^{NL})$. Both of these terms are used to calculate the moment conditions $\{\bar{g}^1(\theta_m), \bar{g}^2(\theta_m), \bar{g}^3(\theta_m)\}$ and, in turn, the objective function that I minimize over the remaining parameters, $\theta_{1m}^{NL} = \{\alpha_m^0, \alpha_m^1, \gamma_m\}$.[27]

I proxy non-grocery expenditure, Z, with household income, Y.[28] To construct the CBSA-month moments, I assume a degenerate distribution for consumer income in each CBSA ($dF(Y|t)$) estimated as a log-normal fitted to the 5-year (2010-2014) average income distribution reported in the ACS for tracts in each CBSA. I therefore identify the non-homotheticity parameters using only household-level purchases as described below.

Identification The market-level moments identify the mean price elasticity, α_m^0 and product quality, β_{mg}, parameters. Conditional on product quality, the base price sensitivity α_m^0 parameter is identified by the extent to which relative within-market sales shares co-vary with the components of relative price variation captured by the price instruments, described in more detail below. Relative product quality, $\tilde{\beta}_{mg} = \beta_{mg} - \beta_{m\bar{g}_m}$, is identified by variation in the average within-market sales shares of each product g, relative to the sales share of the outside product $\bar{g}_m$, conditional on price. The idea here is that, if products with two different products sell at the same price, but product A has a higher average relative market share across all CBSA-months than product B, then product A will be assigned a higher quality parameter relative to the base good for that module.[29]

The household moments identify the non-homotheticity parameters, α_m^1 and γ_m. The α_m^1 parameter that governs how the price sensitivity varies with income is identified primarily by the covariance between the prices households purchase products at and their income. Like α_m^1, the quality-income gradient γ_m parameter that governs how demand for quality varies with income are primarily identified by the covariance between the estimated quality of the products

[27] Details on this full procedure can be found in the documentation provided in the Supplemental Appendix C.

[28] In a slight abuse of notation, I will denote the coefficients on log income using the same notation used to denote the coefficients on log non-grocery expenditure in defining the moment conditions above. These new coefficients are in fact approximations of the original coefficient multiplied by the elasticity of non-grocery expenditure with respect to household income.

[29] Variation in the quality of the outside product across CBSA-months may bias the relative quality estimates that, in practice, are calculated as the mean of CBSA-month-specific quality shocks that rationalize the relative sales shares on that product relative to the outside product given the non-linear parameter estimates, across the CBSA-months in which the product is sold; i.e., $\hat{\tilde{\beta}}_{mg} = \frac{1}{N_g}\sum_{st}\tilde{\beta}_{mgt}(\hat{\theta}_{1m}^{NL})$ where $\tilde{\beta}_{mgt}(\hat{\theta}_{1m}^{NL}) = \beta_{mgt}(\hat{\theta}_{1m}^{NL}) - \beta_{m\bar{g}_m t}(\hat{\theta}_{1m}^{NL})$. I discuss these errors in more detail in Section 6.4.3, where I find them to be small in magnitude and not correlated with the spending patterns of high- or low-income households in such a way that would yield biases in other parameter estimates.

households purchase and their income.

Price Instruments The CBSA-level moments are based on the assumption that $\mathbb{E}[\tilde{\xi}_{mg}(\theta_{1m}^{NL})\tilde{\mathbf{W}}_{mg}^{1}] = 0$ for a set of instruments $\mathbf{W}^1$. These instruments include a set of brand dummies, price instruments, and interaction terms between these sets of variables and moments of the CBSA-level income distribution.[30] These errors and instruments are differenced from the outside product within each market to control, among other things, for market-level variation in the quality of the outside product. The set of brand dummies includes one dummy for each brand except this base product $\bar{g}_m$. To reduce the dimension of the estimation data, I conduct principal components analysis on this final set of instruments and use components that together explain over 95 percent of the variation of the data.[31]

I do not use prices as instruments because they might be correlated with the transient product-market-specific taste shocks, $\xi_{mg}(\theta_{1m}^{NL})$. I instrument for the price charged by stores a given CBSA for a given product with the sales-weighted average contemporaneous price charged for the same product by stores that belong to the set of retail chains as represented in the CBSA but are located in different Demographic Market Areas (geographic market areas defined by Nielsen, which are roughly akin to MSAs). This "same chain-other city" instrument, also employed in DellaVigna and Gentzkow (2019), relies on similar relevance and exogeneity arguments as in Hausman et al. (1994) and Nevo (2001).

For relevance, I rely on cross-product inter-temporal and across-chain variation in the prices charged by chains, driven by the timing of chain-level sales or changes in wholesale pricing arrangements. Recall that the data is differenced from the outside product within market and and implicitly from the product mean, by the inclusion of the product fixed effects. Even after controlling for market and product fixed effects, there is sufficient variation in the instrument to provide a strong first stage, with F-statistics above 30 in all modules and above 150 in 99% of modules.[32]

For exogeneity, cross-product variation in retail chain-level pricing cannot be correlated with changes in relative product tastes in a market. Such a correlation could arise, for example, if prices adjust in response to changes in the tastes of a retail chain's national customer base. A chain might, for example, lower the frequency of promotional sales for a product or re-negotiate a wholesale price agreement in response to declining national demand for that product. Though I am unable to test this exclusion restriction directly, I can – for a subset of my data – construct

[30] Specifically, the average, the average squared, and the standard deviation of the income distribution.

[31] The principal components IV reduces the scale of the optimization problem with minimal sacrifice to identifying variation, noting that linear combinations of valid instruments remain valid instruments – c.f. Bai and Ng (2010). The exact number of principal components used based on Winkelried and Smith (2011)'s retention rule with $\delta = -1.4$. In the typical module, this retains instruments explaining over 98.5 percent of the variation in the instruments, while reducing the number of instruments by 75 percent.

[32] See Appendix Figure A.5.

an instrument that is plausibly uncorrelated with national demand shocks by residualizing my baseline "same chain-other city" instrument from the average contemporaneous price charged for the same product by stores in different DMAs that do not belong to chains represented in the CBSA in question. I use this alternate "other chain-other city" instrument to test the validity of my base instrument in the sub-sample of products over which the residualized instrument is non-missing – i.e., products sold in multiple chains in multiple DMAs.

First, I run a GMM distance test comparing the J-statistics from the model estimated using both "same chain-other city" instruments to the J-statistics from the model estimated using only the residualized version. In most modules, I fail to reject the null that the base instrument is exogenous. Then, I show that the price elasticity estimates using the baseline and the residualized instruments are comparable. Both instruments similarly remove negative biases in the price coefficient relative to an "OLS" specification that uses the endogenous observed price as the instrument (see Appendix Figure A.6). The price coefficients estimated using the base instrument are slightly lower than those estimated using the residualized version, but the difference is small with respective medians of 2.63 and 3.64. In Section 6.4.1 below, I show that the main index results are robust to this increase in the mean price coefficient.

6 Results

6.1 Parameter Estimates

I estimated the model under four sets of parameter restrictions. These restrictions allow preferences to vary with income through the demand elasticities with respect to both quality and price, through only one of these channels, or through neither of these channels, in which case the model is homothetic.

Table 3 summarizes the estimates for the module-level parameters in each of these four models over the 400-550 modules where the optimization procedure reached internal solutions.[33]

Column [1] summarizes the estimates of the parameter that governs the substitution elasticity of a consumer with the mean log income level in the sample for each module, $\hat{\alpha}_m^0 = \hat{\sigma}_m - 1$, for the homothetic version of the model. The median of this price elasticity is 4.2, with an inter-quartile range of 2.4 to 6.5. Allowing for non-homothetic demand for quality and/or price in columns [2], [4], and [6], the median price elasticity falls to between 2 and 2.6 (implying a median elasticity of substitution between 3 and 3.6). These own-price elasticities are in the range of those estimated for similar categoreis of products in Nevo (2000), Dube (2004), and Faber and Fally (2017).

[33]The parameters were bounded as follows: $\alpha_m^0 \in (0.05, 30)$, $\alpha_m^1 \in (-5, 5)$, and $\gamma_m \in (-5, 5)$.

Table 3: Summary Statistics for Parameter Estimates

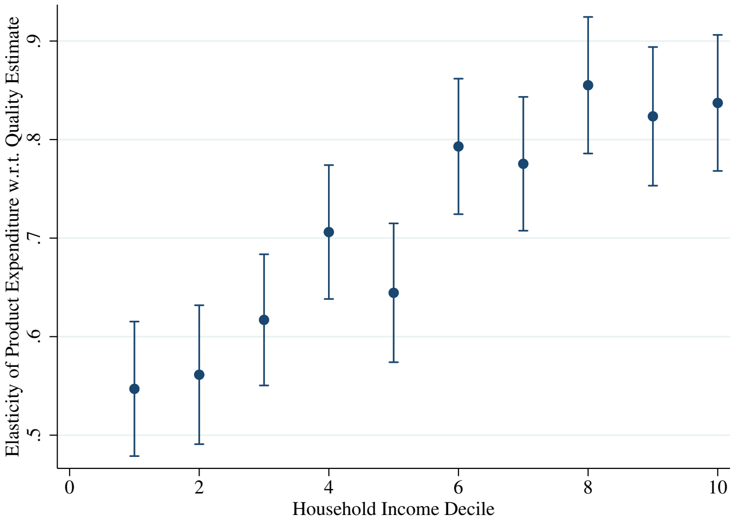
Model:	Homothetic	NH in Quality		NH in Price		NH in Quality and Price		
Restrictions:	$\alpha_m^1 = 0$ & $\gamma_m = 0$	$\alpha_m^1 = 0$		$\gamma_m = 0$		None		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Parameter:	α_m^0	α_m^0	γ_m	α_m^0	α_m^1	α_m^0	α_m^1	γ_m
Count	438	516	516	491	491	555	555	555
with t>1.96	392	494	476	482	372	453	71	462
with t<-1.96	0	0	27	0	104	0	351	31
Mean	5.50	4.29	1.22	3.43	1.46	2.79	-1.52	1.74
p25	2.44	1.73	0.56	1.03	0.58	1.41	-2.48	1.01
p50	4.21	2.63	1.00	1.95	1.28	2.44	-1.36	1.58
p75	6.53	5.21	1.64	3.56	2.50	3.59	-0.62	2.50

Notes: These tables report the summary statistics for the main module-level parameter estimates governing the elasticity of substitution and non-homotheticities in demand. Attention is limited to modules for which the estimation procedure converged at interior estimates for all relevant parameters. The second and third rows of the table show the number of modules in which the estimated t-statistic for the parameter was above or below 1.96. The mean and percentile statistics in the subsequent rows are weighted by module sales in the Nielsen store-level data. The full distributions of the γ_m and α_m^1 estimates are depicted in Supplemental Appendix Figure (B.2).

Columns [3] and [8] of Table 3 summarize the distribution of the estimated values for γ_m. All four models assume that all consumers agree on the relative quality of products, as described by the distribution of the β_{mg} parameters for products $g \in \mathbf{G_m}$ within a module m. For positive values of γ_m, however, the utility weight that consumers place on this component of utility, relative to their idiosyncratic utility draw for each product or the quantity consumed, is increasing in their non-grocery expenditure Z. This implies that consumers with higher non-grocery expenditures have a higher willingness to pay for quality. In estimation, these parameters are identified by the fact that higher income consumers spend a relatively greater share of module expenditure on products with relatively high β_{mg} estimates, that is, the products for which all consumers have a higher willingness to pay. Figure 4 shows that products with higher β_{mg} estimates have higher expenditures at all income levels, but more so for the rich. Accordingly, Columns [3] and [8] of Table 3 show that the willingness to pay for quality (governed by γ_m) increases with income in over three-quarters of the modules represented in the data. The demand for quality is therefore increasing with income in most grocery sectors.

Columns [5] and [7] of Table 3 summarize the distribution of the estimated values for α_m^1 in each module. Recall this parameter governs how the elasticity of substitution varies across consumers with different non-grocery expenditures. For $\alpha_m^1 < 0$, high-income consumers will find other products to be less substitutable with their ideal variety and, therefore, substitute less across products in response to relative price changes. Comparing columns [5] and [7] of Table 3, we see that the majority of the α_m^1 estimates are instead positive unless you control for non-homotheticity in the demand for quality. Column [5] shows that the majority of the α_m^1

Figure 4: Product Quality (β_{mg}) Estimates and High-vs.-Low Income Household Expenditures



Note: Plots shows coefficient on log product-level expenditures by each income decile in the household-level (HMS) data regressed against the product quality (β_{mg}) estimates in the model that allows for non-homotheticity in quality but not price sensitivity (i.e., restricting $\alpha^1_m = 0$ but allowing $\gamma_m \neq 0$). These regressions include product module fixed effects and observations are weighted by aggregate module sales. Attention is limited to estimates in the modules where the estimation procedure converged at interior estimates.

estimates, and even the majority of those that are statistically significant, are instead positive when γ_m is constrained to be zero.[34] Column [7], on the other hand, shows that, in over 75 percent of modules, high-income consumers are less price sensitive, or $\hat{\alpha}^1_m < 0$, when you control for the fact that they also have a greater willingness to pay for quality.

The parameter estimates generally support that demand is non-homothetic within modules. In particular, high-income consumers have a greater willingness to pay for quality than low-income consumers and, when controlling for this non-homotheticity in the demand for quality, the results show that high-income consumers are also less price sensitive.[35]

[34]These estimates may be biased upwards by a correlation between unobserved income-specific product tastes and prices. Consider the model where γ_m is restricted to equal zero for a degenerate CBSA income distribution: $\ln s_{mgt} - \ln s_{m\bar{g}_m t} = (\alpha^0_m + \alpha^1_m y_{st})[(\beta_{mg} - \beta_{m\bar{g}_m}) - (\ln p_{mgt} - \ln p_{m\bar{g}_m t})] + \nu_{mg\bar{g}_m t}$. If the true γ_m is positive, the error terms here will include any income-specific product tastes, $\gamma_m\left(\beta_{mg} - \beta_{m\bar{g}_m}\right)$. If stores in high-income CBSAs set prices in accordance with these tastes such that $Corr(\gamma_m\left(\beta_{mg} - \beta_{m\bar{g}_m}\right), \ln p_{mgt} - \ln p_{m\bar{g}_m t}) \neq 0$, then the assumption that $\mathbb{E}[\mathbf{W}\xi] = 0$ will be violated. The fact that the α^1_m estimates are lower, and generally negative, in the model that allows for non-homotheticity in the demand for quality and the price sensitivity supports this theory, since this model directly controls for, $\gamma_m y_t(\beta_{mg} - \beta_{m\bar{g}_m})$. I do not, therefore, take the positive α^1_m estimates in the model that does not control for correlations in income-product specific tastes as evidence that high-income consumers are more price sensitive than low-income consumers. Instead, the positive α^1_m estimates highlight the difficulty in identifying the non-homotheticity related to price sensitivity in isolation from the non-homotheticity related to product quality.

[35]Supplemental Appendix D.2 provides further evidence with moments demonstrating the out-of-sample fit of the model.

6.2 Model Selection

The model estimates above provide micro-evidence that high-income households have a stronger taste for high-quality products and, controlling for this, they are less price sensitive. Allowing for both forms of non-homotheticity introduces around 500 additional parameters to the model (one α^1_m or γ_m for each module). These parameters will all be sources of error in the income-specific price indexes used to address the paper's main question in Section 6.3 below. Prior to undertaking this analysis, I therefore first attempt to determine whether this parametric flexibility is valuable enough to warrant these additional errors. To do this, I use the GMM-BIC model selection criterion that judges models using a trade-off between model fit and model complexity, measured using the number of parameters relative to the number of moments used in the estimation of those parameters. Specifically, for each module, the GMM-BIC criterion selects the model and moment conditions that minimize the difference between the estimated J statistic and the log of the number of observations multiplied by the number of over-identifying restrictions used in estimation.[36]

The model that permits non-homothetic demand for quality, but not for price, dominates the models that permit non-homothetic demand for price or both price and quality in over 80 percent of modules, representing 81 and 88 percent of sales, repectively. The model that accounts for non-homothetic demand for quality has a lower GMM-BIC criterion than both of the alternative non-homothetic models in over 70 percent of modules, representing 74 percent of sales.

These results suggest that the salient form of non-homotheticity in grocery consumption is in the demand for quality. In the analysis below, I limit my attention to price indexes that account for this form of non-homotheticity alone when studying how grocery costs vary across local markets differently for consumers at different income levels. Any differences between the relative price indexes high- and low-income consumers face across cities and stores will reflect differences in the availability and prices of high- relative to low-quality products across these markets.[37]

[36]This method was developed in Andrews (1999) as a moment selection criterion and is shown to be consistent for model selection in Andrews and Lu (2001). The selection criterion minimizes the following GMM-BIC function:

$$\text{GMM-BIC}^M_m(\hat{\theta}^M_{1m}) = n_m G_m(\hat{\theta}^M_{1m}, \bar{\theta}^M_{1m})' W^*_m G_m(\hat{\theta}^M_{1m}, \bar{\theta}^M_{1m}) - \ln(n_m)(L^*_m - K^M_m) \tag{19}$$

where $G_m(\hat{\theta}^M_{1m}, \bar{\theta}^M_{1m})$ are the moments for model M evaluated at the estimated values for free parameters $\hat{\theta}^M_{M1m}$ and zero for the restricted parameters, $\bar{\theta}^M_{1m}$; K^M_m is the number of free parameters in model M for module m; and n_m and L^*_m are the number of observations and instruments, respectively, used to estimate all models for module m. The same set of instruments is used to calculate each moment condition, and thus the number of moments is also common between models for each module. W^*_m is the optimal weighting matrix for the full model.

[37]Conversely, these price indexes do not allow for non-homotheticity in consumer's price sensitivity (or idiosyncratic utility weight). So, while high-income consumers face relatively lower costs in markets with relatively more, and cheaper, high-quality products than low-quality products, all consumers get the same additional utility, and cost savings, in markets that offer more varieties and lower prices of both high- and low-quality products

6.3 Income-Specific Consumption Externalities

The analysis above has provided the inputs to market- and income-specific price indexes that represent how households at different income levels value the products and prices available to them in different U.S. cities and neighborhoods, as outlined in Section 5 above. I can now turn to answering the central question in this paper: do grocery costs vary differently across markets for consumers at different income levels?

To answer this question, I estimate the following regression:

$$\ln \hat{P}(\mathbb{P}_c, y_k) = \delta_k + \beta_1 y_c + \beta_2 (y_k - \bar{y}_k) y_c + \epsilon_{kc}, \tag{20}$$

where $\hat{P}(\mathbb{P}_c, y_k)$ is the grocery price index for a representative consumer with log income y_k in each market c, obtained by plugging the market-specific price vector $\mathbb{P}_c$, income y_k, and model parameter estimates into equation (15); δ_k is an income-level fixed effect; y_c is log per capita income in city c, and $\bar{y}_k$ is the mean log household income in the sample.[38]

The coefficient on log city income (β_1) reflects the mean elasticity of grocery costs with respect to city income. The coefficient on the interaction of demeaned log consumer income and log city income (β_2) measures how the elasticity varies with household income. The grocery price index, $\hat{P}(\mathbb{P}_c, y_k)$, is calculated using a model that allows for non-homotheticity in the demand for quality, so the elasticity of grocery costs with respect to city income will vary with income, and β_2 will be non-zero, if the goods and prices available in each city are correlated with the tastes corresponding to the average income of the consumers living there. If wealthy cities offer more varieties of high-quality goods at lower prices than poorer cities, the price index faced by high-income consumers will decrease by more (or increase by less) than the price index faced by low-income consumers between poor and wealthy cities. This is because high-income consumers benefit more from the availability and lower prices of the goods that they prefer. Under this scenario, the elasticity of the price index faced by high-income consumers with respect to city income would be lower than the elasticity of the price index faced by low-income consumers with respect to city income yielding a negative β_2 estimate.[39]

Table 4 presents the results of the baseline regression estimated using income-specific price indexes calculated for price vectors reflecting the prices and products available at 100 random

equally.

[38] In practice, the quality of the base product in each module ($\beta_{m\bar{g}_m}$) is not identified in estimation, so the relative product qualities ($\tilde{\beta}_{mg} = \beta_{mg} - \beta_{m\bar{g}_m}$) are used in place of the absolute product qualities to calculate $\hat{P}(\mathbb{P}_c, y_k) = P_m^{CES}(\mathbb{P}_c, y_k)/\mathbb{B}(y_k)$, where $\mathbb{B}(y_k) = \prod_{m \in \mathbf{M}} \left(\exp(\beta_{m\bar{g}_m} \gamma_m(y_k))\right)^{\lambda_m}$ is a residual market-invariant base-quality aggregator that is controlled for with the income-level fixed effect, δ_k.

[39] This regression characterizes an equilibrium relationship and should not be interpreted causally. The results presented here do not indicate whether, for example, grocery costs are lower for high-income consumers in wealthy cities because a high per capita income causes stores in a city to stock more high-quality products or because high-quality products attract more high-income inhabitants to a city, raising its per capita income.

samples of 50 stores in each of the 125 CBSAs that have 50 or more stores.[40] The β_1 coefficient on log CBSA per capita income is negative but not significant, reflecting the large degree of noise in the price indexes across CBSAs making it impossible to identify a systematic relationship between the mean price index that a household faces in a city and its per capita income. There is, on the other hand, strong evidence that the elasticity of the price index with respect to per capita income increases with household income: the β_2 coefficients on the interaction between log CBSA per capita income and demeaned log household income are negative and statistically significant. The magnitude of the β_2 estimate indicates that this variation is economically significant. A consumer who earns $25,000 a year sees their per dollar grocery costs increase by around 14 percent for each log unit increase in city per capita income, comparable to the gap between the wealthiest and poorest cities in the sample (Bridgeport-Stamford-Norwalk, CT with per capita income of $49,688 and El Paso, TX with per capita income of $18,684). On the other hand, the per dollar grocery costs of a consumer with a yearly income of $200,000 decrease by 26 percent for each log unit increase in city per capita income. A high-income household would experience an 7 percent greater decrease in grocery costs than a low-income household when both move from a CBSA at the 25th percentile of the income distribution (e.g., San Antonio, TX) to a CBSA at the 75th percentile of the income distribution (e.g., Providence, RI).

Market income is correlated with market size: in this sample, wealthier cities are larger than poorer cities with a correlation coefficient of 0.35. Therefore, it is possible that a negative β_2 estimate in the baseline regression could result from grocery costs being lower for high-income households than for low-income households in larger, as opposed to wealthier, cities. In column [2] of Table 4, I therefore add controls for log population and log population interacted with log household income to the baseline regression. The β_2 coefficient is robust to these controls, whose coefficients are estimated as precise zeros. This evidence is consistent with the "within-group preference externalities" story in which higher income consumers receive relatively more consumption benefits from living in wealthier cities, as opposed to a story in which high-income consumers receive more consumption benefits from living in larger cities than low-income consumers.

[40]Formally, the regression estimated is:

$$\ln \hat{P}(\mathbb{P}_{cb}, y_k) = \delta_{kb} + \beta_1 y_c + \beta_2 (y_k - \tilde{y}_k) y_c + \epsilon_{kcb},$$

where $\mathbb{P}_{cb}$ denotes the set of prices available to consumers in the 50 stores in bootstrap sample b for CBSA c and δ_{kb} is a bootstrap sample-household income group fixed effect. Standard errors are clustered at the CBSA level. This regression estimates log-linear relationships between CBSA income and the semi-elasticity of the price level with respect to household income and between household income and the semi-elasticity of the price level with respect to CBSA income. In Supplemental Appendix D.3, I estimate these relationships non-parametrically and find them to be close to log-linear.

Table 4: City-Income Specific Price Index Regressions

Dependent Variable: Ln(Price Index for Household in Income Group k in CBSA c)	Local Prices		National Prices	
	[1]	[2]	[3]	[4]
Ln(Per Capita Income$_c$) (β_1)	-0.068	-0.042	-0.18*	-0.17
	(0.088)	(0.10)	(0.084)	(0.095)
Ln(Per Capita Income$_c$)* Demeaned Ln(HH Income$_k$) (β_2)	-0.18***	-0.15***	-0.21***	-0.19***
	(0.038)	(0.039)	(0.042)	(0.044)
Ln(Population$_c$) (β_3)		-0.0095		-0.0052
		(0.018)		(0.018)
Ln(Population$_c$)* Demeaned Ln(HH Income$_k$) (β_4)		-0.011		-0.0077
		(0.0072)		(0.0069)
Income Group k*Bootstrap Sample FEs	Yes	Yes	Yes	Yes
Number of CBSAs (c)	125	125	125	125
Observations	100,000	100,000	100,000	100,000
adj. within R^2	0.02	0.02	0.05	0.05

Notes: *** p< 0.01, ** p<0.05, * p<0.10; standard errors, clustered by bootstrap sample and CBSA, are in parentheses. This table presents results from regressions of household income- and CBSA-specific grocery price indexes against CBSA characteristics alone and interacted with demeaned log household income. The price indexes correspond to the baseline model that allows for non-homotheticity in the demand for quality but not in price sensitivity (i.e., restricting that α_m^1=0) and measure how households at eight different income levels between \$25,000 and \$200,000 value the products and prices represented in each of 100 bootstrap samples of 50 stores in each of 125 CBSAs with 50 or more participating retailers.

Differentiating between Price and Variety Effects

The results above suggest that, relative to low-income households, high-income households receive higher consumption utility from the grocery bundles available in wealthier cities than from the grocery bundles available in poorer cities with the same population. The model allows for high-income households to have a stronger preference for high-quality goods than do low-income households. So, the fact that high-income households get relatively more utility from consuming grocery products in high-income cities must be either because there are more high-quality goods available in these locations or because the high-quality goods are sold at relatively lower prices in high-income cities, or for both reasons. I examine this issue by calculating income-specific price indexes for the set of products I observe in the 50-store sample for each city, as before, but setting the prices of each product equal to its national average price.

Columns [3] through [4] of Table 4 replicate columns [1] through [2] using these fixed-price indexes as the dependent variable. The coefficients on the interaction between per capita income and household income increase slightly in magnitude, but the change is not statistically significant. High-income households would continue to find wealthy cities almost as cheap relative to poor cities, relative to low-income households, if products were sold in both locations at their national average price. This indicates that the difference in how high- and low-income households perceive the relative costs to vary across cities is due to variety differences. Prices are higher in wealthy cities relative to poor cities, but high-income consumers are more than compensated for this price difference by the fact that more of the products they prefer to consume are available to them in these locations.[41]

Variation within CBSAs

We see similar variation in the per dollar grocery utility offered to high- and low-income households across stores in different neighborhoods as we did across CBSAs. Table 5 presents the elasticity estimates from equation (20) where market s denotes a store located in CBSA $c(s)$.[42] Column [1] shows a similar pattern in the variation in the elasticity of price indexes with respect to household income across stores with different local per capita income as we saw across CBSAs with different per capita income. With these store-level indexes, we can consider whether sorting within CBSAs might enable households to mitigate some of the cross-CBSA variation in grocery availability. Column [2] shows that the elasticity of store-level indexes with respect

[41] Appendix Figure A.4 shows that wealthy cities offer more variety but also charge higher prices. For most households (earning \$100,000 or less), however, the greater variety offered in wealthy cities is insufficient compensation for the higher prices.

[42] For the store-level results, $\hat{P}(\mathbb{P}_s, y_k)$ reflects the grocery price index of a representative household earning y_k faces in store s and y_s is the average size-adjusted income in the vicinity of store s, calculated using the non-parametric method described in Appendix A.2..

to household income is also increasing with CBSA income. Columns [3] and [4] show that this correlation is stronger when comparing the indexes for stores located in the high-income neighborhoods in different CBSAs. That is, the relationship between grocery costs and CBSA income is amplified for residents of high-income neighborhoods and mitigated for residents of low-income neighborhoods.

Table 5: Store Price Index Regressions

Dependent Variable: Ln(Price Index for Representative Consumer k in Store s in CBSA $c(s)$)	[1]	[2]	[3]	[4]	[5]	[6]
Ln(Per Capita Income$_s$) (β_1)	-0.097***				0.058***	
	(0.0057)				(0.0040)	
Ln(Per Capita Income$_s$)* Demeaned Ln(HH Income$_k$) (β_2)	-0.20***				-0.018***	
	(0.0050)				(0.0023)	
Ln(Per Capita Inc$_c(s)$) (β_3)		-0.13***	-0.18***	-0.013		0.037***
		(0.024)	(0.045)	(0.062)		(0.010)
Ln(Per Capita Income$_c(s)$)* Demeaned Ln(HH Income$_k$) (β_4)		-0.21***	-0.17***	-0.091*		-0.020*
		(0.022)	(0.045)	(0.044)		(0.0098)
Income Group k FEs	Yes	Yes	Yes	Yes	Yes	Yes
Chain x Income Group FEs	No	No	No	No	Yes	Yes
Store Set (Local Per Capita Income$_c$)	All	All	High-Inc.	Low-Inc.	All	All
Number of Stores (c)	9330	8894	4653	4241	9329	8893
Number of CBSAs	-	689	172	649	-	689
Observations	74,640	71,152	37,224	33,928	74,632	71,144
adj. within R^2	0.06	0.06	0.05	0.00	0.01	0.00

Notes: *** p< 0.01, ** p<0.05, * p<0.10; standard errors, clustered by store and household income in columns 1 and 5 and by CBSA and household income in columns 2 through 4 and 6, are in parentheses. This table presents results from regressions of household income- and store-specific grocery price indexes against measures of local store income alone and interacted with demeaned log household income. The price indexes correspond to the baseline model that allows for non-homotheticity in the demand for quality but not in price sensitivity (i.e., restricting that α_m^1=0) and measure how households at eight different income levels between $25,000 and $200,000 value the products and prices represented in grocery stores in the Nielsen RMS sample. Store-by-income group observations are weighted by store sales.

The results in columns [5] and [6] show that the variation in columns [1] through [2] is almost entirely explained by variation in the set of retail chains that locate in high- vs. low-income neighborhoods. Retail chains do not appear to significantly alter the mix of brands they offer across neighborhoods or CBSAs in a way that biases the attractiveness of their stores in higher-income locations to higher-income customers.

6.4 Robustness Checks

6.4.1 Robustness to Different Estimation Choices

Table 6 shows the robustness of the demand parameters and index elasticities estimated above to various decisions made in the course of estimation. Due to computation limitations, the main estimation procedure grouped any products with expenditure shares below the 60th percentile in a given CBSA-month to an outside product for that CBSA-month and then drops any CBSA-month markets where this outside product accounts for less than 3 percent of sales. The first column replicates the median key parameter values and index elasticities under this base specification. The next three columns show the robustness of key parameter estimates to allocating either fewer or more products (those below the 40th or 80th percentiles) to the outside product and to dropping markets where the outside product accounts for less than 1 (rather than 3) percent of sales. The next column shows the results when the estimation data are aggregated to the quarterly, instead of monthly, frequency, and the final column shows the results from the specification employing the residualized instrument described in Section 5.3.1. The first two rows show the median price elasticity (α_m^0) and income-quality gradient (γ_m) estimates, while subsequent rows replicate the main specification from Table 4 for price indexes calculated using the parameter estimates from each of these robustness specifications.

Reassuringly, the parameter estimates and index elasticites are relatively stable. There is of course some variation in the parameter estimates across specifications. The median price elasticity (α_m^0) estimates (in column [2]) range between 2.3 and 3, and increase to 3.6 with the residualized instrument, while the median estimates for income-quality gradient (γ_m) estimates (in column [5]) fall between 0.66 and 1.20 across all specifications. The relative stability of the γ_m estimates, in particular, translates in to rather stable estimates for the cross-elasticity of the associated price indexes with respect to city and household income in Table 6. This cross-elasticity varies between -0.11 and -0.31, with the lowest elasticities in the specifications that yield the lowest income-quality gradient (γ_m) estimates. The estimate for the base specification falls in the middle of this band. Together, these results confirm that high-income households find wealthy cities less expensive than poor cities relative to low income households.

6.4.2 Outlier Modules

One might be concerned that the results above are driven by a small number of product categories with outlier demand parameter estimates. To study the role of outliers, I replicate the regression in column [1] of Table 4 module-by-module. The sales-weighted distribution of the resulting module-level coefficients on the per capita income-household income interaction term is clustered between -0.5 and 0. There are a few outliers, but these product categories reflect

Table 6: Robustness of Index Elasticities to Alternative Specifications

Dependent Variable: Ln(Price Index for Household in Income Group k in CBSA c)

Estimation Specification:	Base	OG 40%	OG 80%	OG Sh > 1%	Qtly Data	Resid IV
	[1]	[2]	[3]	[4]	[5]	[6]
Median Price Elasticity (α_0)	2.63	2.90	2.36	2.72	2.51	3.64
Median Income-Quality Elasticity (γ)	1.00	1.06	1.20	1.19	0.87	0.66
Ln(Per Capita Income$_c$) (β_1)	-0.068	-0.16	-0.092	-0.063	-0.035	-0.015
	(0.088)	(0.12)	(0.11)	(0.091)	(0.080)	(0.060)
Ln(Per Capita Income$_c$)*	-0.18^{***}	-0.31^{***}	-0.18^{**}	-0.21^{***}	-0.14^{***}	-0.11^{***}
Demeaned Ln(HH Income$_k$) (β_2)	(0.038)	(0.039)	(0.062)	(0.046)	(0.028)	(0.016)
Income Group k*Bootstrap Sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Number of CBSAs (c)	125	125	125	125	125	125
Observations	100,000	100,000	100,000	100,000	100,000	100,000
adj. within R^2	0.02	0.03	0.01	0.02	0.01	0.01

Notes: *** p< 0.01, ** p<0.05, * p<0.10; standard errors, clustered by bootstrap sample and CBSA, are in parentheses. The first two rows of this table present the median price elasticity and income-quality gradient estimates obtained in the baseline as well as various robustness specifications. The subsequent rows present results from regressions of household income- and CBSA-specific grocery price indexes calculated using parameter estimates from each of these specifications against log CBSA per capita income alone and interacted with demeaned log household income. The parameter estimates and price indexes are for the baseline model that allows for non-homotheticity in the demand for quality but not in price sensitivity (i.e., restricting that α^1_m=0). The price indexes measure how households at eight different income levels between \$25,000 and \$200,000 value the products and prices represented in each of 100 bootstrap samples of 50 stores in each of 125 CBSAs with 50 or more participating retailers.

only a small share of sales so, under the Cobb-Douglas demand assumption, cannot drive the cross-elasticity of the aggregate price indexes.

6.4.3 Measurement Error in Quality

To estimate product quality, I have assumed that the quality of the outside good in each module is equal across markets. In practice, variation in the quality of the outside product across store-months will generate errors in the relative quality estimates ($\tilde{\beta}_{mg}$). One concern is that quality may be mis-measured in a way that biases the gradient of the quality elasticity with respect to income (γ_m). For example, suppose that high-income households tend to purchase products in markets that also offer higher quality outside goods. $\tilde{\beta}_{mg}$ will then understate the relative quality of products that high-income households purchase, and overstate the relative quality of products that low-income households purchase. This could lead me to overstate the income-quality elasticity gradient (γ_m).[43]

[43] Alternatively, if the bias is so large that the ordering of product quality is not maintained – such that products that high-income households favor are estimated to have lower relative quality than products low-income households favor when they are in fact higher quality (or vice versa) – I could estimate the wrong sign for the income-quality elasticity gradient (γ_m). In this case, the main result that high-income markets offer more of the products that high-income households favor and, therefore, provide high-income households with relatively lower grocery costs than low-income markets, would hold, but the interpretation that these products are higher quality (i.e., preferred on average by all households) would not.

I run two tests to gauge the degree of this error and its potential to bias the γ_m estimates. The results, in Appendix D.2, show that these errors are typically small in magnitude. Importantly, I find that the errors are not correlated with the purchasing behavior of high- vs. low-income households in such a way that would bias the income-quality elasticity gradient. The robustness of the γ_m and β_{mg} parameter estimates to alternate definitions of the outside good in Table 6 is also reassuring.

6.4.4 Alternative Sources of Demand Heterogeneity

The price indexes calculated here account for how consumer tastes vary with income both across products in the same category and across categories of products. Income is a factor in determining a consumer's preferences over different types of breakfast cereal, for example, as well as in determining their willingness to pay for cereal relative to milk. In order to make this multi-sector analysis tractable, I have abstracted from a number of other ways in which demand and, therefore, aggregate costs could vary across heterogeneous households.

In particular, empirical micro-economists have shown that income is just one of a range of demographic characteristics that can be correlated with consumer demand for a variety of product characteristics, including brand quality. The model here is more stylized, allowing the willingness to pay for a single product characteristic, brand name, to vary with a single consumer characteristic, income. The benefit of such a simple framework is that it is generalizable: none of the variables are category-specific so it can be used to measure how demand varies systematically with consumer characteristics across products in many product categories. The drawback is that it imposes two types of strong assumptions on the consumer tastes.

The first is that households value units of products from the same brand and module equally, regardless of their flavor, texture, or the size and type of container they were packaged in. The cross-city price indexes I calculate account for the fact that high-quality brand name products are more available or sold at cheaper prices than low-quality brand name products in some cities than in others, but the prices of products in the same module and brand enter symmetrically, even if they have different sizes, container types, etc.. For violations of this assumption to bias the results of the paper, low-income tastes would need to be biased towards product characteristics that are disproportionately represented (or available at lower prices) in high-income cities. This is unlikely to be the case. I do not, for example, find any statistically significant correlations between either the price or availability of products with certain sizes and per capita income when controlling for product module and brand name.

The second simplification in the model above is that, controlling for size-adjusted household income, consumer demand does not vary systematically with other demographics, such as age, marital status, and education. The consumption patterns and parameter estimates above are consistent with non-homotheticicities in demand but may instead pick up correlations between demand and these other demographics, to the extent that age, marital status, and education are also correlates of income. Similarly, the estimated patterns in product availability across high- and low-income markets are consistent with local firms catering to income-specific tastes, but could also be the result of preference externalities along other demographics or unrelated supply-side factors. It is important, therefore, to caution against interpreting these results causally. More work is needed to assess the role of preference externalities in grocery retail.

7 Conclusion

There is growing interest in the role of non-homothetic preferences and cross-market income differences in determining production patterns in macro, urban, and international economics. If preferences are income-specific and, further, if the products available in different markets are biased to the income-specific tastes in these markets, then consumers at different income levels will experience different changes in consumption utility across these markets. The results in this paper indicate that this is indeed the case: high-income households face greater grocery consumption gains from moving to high per capita income markets than do low-income households.

I show that high-income households face much lower grocery costs in wealthy cities than in poor cities, while low-income households face slightly higher grocery costs in these locations. Further work is required to extend the analysis presented here to other components of household expenditure in order to build income-specific aggregate spatial price indexes that can be used, for example, in real income measurement or in a Rosen-Roback framework to look at the role of these pecuniary consumption amenities, relative to skill-biased productivity spillovers, in explaining skill-biased agglomeration. Recent work by Atkin et al. (2020) suggests a promising path forward in this direction.

I do not expect that these grocery cost differentials are representative of the differentials that we would expect in other components of the typical consumer basket. For one, I expect that the availability of the food and fast-moving consumer goods represented in my sample varies less geographically than other parts of the consumption basket like non-tradable services and housing. If anything, I would expect the strength of consumption externalities to be higher in sectors that are less tradable. So, conditional on these other products having similar degrees of demand heterogeneity, I would consider my estimates to be a lower bound for the differentials we would expect to see in aggregate price indexes.

References

Adams, B. and K. R. Williams (2019). Zone pricing in retail oligopoly. *American Economic Journal: Microeconomics 11*(1), 124–56.

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

Albouy, D. (2009, August). The Unequal Geographic Burden of Federal Taxation. *Journal of Political Economy 117*(4), 635–667.

Albouy, D., G. Ehrlich, and Y. Liu (2016). Housing demand, cost-of-living inequality, and the affordability crisis. NBER Working Paper No. 22816.

Almas, I. (2012). International Income Inequality: Measuring PPP Bias by Estimating Engel Curves for Food. *The American Economic Review 102*(2), 1093–1117.

Anderson, S. P., A. de Palma, and J.-F. Thisse (1987). The CES is a Discrete Choice Model? *Economics Letters 24*, 139–140.

Andrews, D. W. K. (1999, May). Consistent Moment Selection Procedures for Generalized Method of Moments Estimation. *Econometrica 67*(3), 543–564.

Andrews, D. W. K. and B. Lu (2001). Consistent Model and Moment Selection Procedures for GMM Estimation with Application to Dynamic Panel Data Models. *Journal of Econometrics 101*, 123–164.

Argente, D. and M. Lee (2016). Cost of living inequality during the great recession. *Journal of the European Economic Association*.

Atkin, D., B. Faber, T. Fally, and M. Gonzalez-Navarro (2020). Measuring welfare and inequality with incomplete price information. NBER Working Paper No. 26890.

Auer, R. (2010, August). Consumer Heterogeneity and the Impact of Trade Liberalization: How Representative is the Representative Agent Framework? *Swiss National Bank Working Paper* (13).

Bai, J. and S. Ng (2010). Instrumental Variable Estimation in a Data Rich Environment. *Econometric Theory 26*, 1577–1606.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Berry, S., J. A. Levinsohn, and A. Pakes (2004). Differentiated Products Demand Systems from a Combination of Micro and Macro Data: the New Car Market. *Journal of Political Economy 112*(1), 68–105.

Bils, M. and P. J. Klenow (2001). Quantifying Quality Growth. *The American Economic Review 91*(4), 1006–1030.

Black, D., N. Kolesnikova, and L. Taylor (2009). Earnings functions when wages and prices vary by location. *Journal of Labor Economics 27*(1), 21–47.

Broda, C., E. Leibtag, and D. E. Weinstein (2009). The role of prices in measuring the poor's living standards. *The Journal of Economic Perspectives 23*(2), 77–97.

Broda, C. and J. Romalis (2009). The Welfare Implications of Rising Price Dispersion. NBER Working Paper No. 18314.

Cage, R. A., T. Garner, and J. Ruiz-Castillo (2002). Constructing household specific consumer price indexes: An analysis of different techniques and methods. BLS Working Paper No. 354.

Couture, V. and J. Handbury (2020). Urban revival in America. *Journal of Urban Economics 119*, 103–267.

Deaton, A. (2010). Price indexes, inequality, and the measurement of world poverty. *American Economic Review 100*(1), 5–34.

Deaton, A. and O. Dupriez (2011a, April). Purchasing Power Parity Exchange Rates for the Global Poor. *American Economic Journal: Applied Economics 3*(2), 137–166.

Deaton, A. and O. Dupriez (2011b). Spatial Price Differences within Large Countries. *Woodrow Wilson School of Public and International Affairs Working Paper* (1321).

Deaton, A. and J. Muellbauer (1980). *Economics and Consumer Behavior*. Cambridge University Press.

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

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review 106*(3), 479–524.

Dingel, J. I. (2016). The determinants of quality specialization. *The Review of Economic Studies 84*(4), 1551–1582.

Dube, J.-P. (2004). Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks. *Marketing Science 23*(1), 66–81.

Faber, B. (2014). Trade liberalization, the price of quality, and inequality: Evidence from mexican store prices. *UC-Berkeley Working Paper*.

Faber, B. and T. Fally (2017). Firm heterogeneity in consumption baskets: Evidence from home and store scanner data. NBER Working Paper No. 23101.

Fajgelbaum, P., G. M. Grossman, and E. Helpman (2011, August). Income Distribution, Product Quality, and International Trade. *Journal of Political Economy 119*(4), 721–765.

Feenstra, R. C. (1994). New Product Varieties and the Measurement of International Prices. *The American Economic Review 84*(1), 157–177.

Feenstra, R. C. and J. Romalis (2014). International prices and endogenous quality. *The Quarterly Journal of Economics 129*(2), 477–527.

Garner, T. I., D. S. Johnson, and M. F. Kokoski (1996). An experimental consumer price index for the poor. *Monthly Lab. Rev. 119*, 32.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer city. *Journal of Economic Geography 1*, 27–50.

Hallak, J. C. (2006). Product Quality and the Direction of Trade. *Journal of International Economics 68*(1), 238–265.

Handbury, J. and D. E. Weinstein (2014). Goods prices and availability in cities. *The Review of Economic Studies 82*(1), 258–296.

Hausman, J. A., G. Leonard, and J. D. Zona (1994). Competitive Analysis with Differentiated Products. *Annals D'Economie et Statistique 34*, 159–180.

Hendel, I. (1999). Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns. *The Review of Economic Studies 66*(2), 423–446.

Hitsch, G. J., A. Hortaçsu, and X. Lin (2019). Prices and Promotions in US Retail Markets: Evidence from Big Data. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2019-117).

Hortaçsu, A. and J. Joo (2015). Semiparametric estimation of a CES demand system with observed and unobserved product characteristics. *arXiv preprint arXiv:1511.05600*.

Hottman, C. (2014). Retail markups, misallocation, and store variety in the US. *Mimeograph, Columbia University*.

Hummels, D. and V. Lugovskyy (2009). International Pricing in a Generalized Model of Ideal Variety. *Journal of Money, Credit and Banking 41*(1), 3–33.

Jaravel, X. (2018). The unequal gains from product innovations: Evidence from the us retail sector. *The Quarterly Journal of Economics 134*(2), 715–783.

Jorgenson, D., D. T. Slesnick, and D. Slottje (1989). Redistributional policy and the measurement of poverty. *Research on Economic Inequality*.

Kokoski, M. F. (1987). Consumer price indexes by demographic group. *Bureau of Labor Statistics (Washington DC) Working Paper* (167).

Li, N. (2021). An Engel curve for variety. *Review of Economics and Statistics 103*(1), 72–87.

Manson, S., J. Schroeder, D. Van Riper, and S. Ruggles (2018). IPUMS National Historical Geographic Information System: Version 13.0.

Moretti, E. (2013). Real Wage Inequality. *American Economic Journal: Applied Economics 5*(1), 65–103.

Nevo, A. (2000). Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry. *The RAND Journal of Economics 31*(3), 395–421.

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica 69*(2), 307–342.

Pinjari, A. R. and C. Bhat (2010). A multiple discrete-continuous nested extreme value (MDC-NEV) model: Formulatin and application to non-worker activity time-use and timing behavior on weekdays. *Transportation Research Part B: Methodological 44*(4), 562–583.

Simonovska, I. (2015). Income differences and prices of tradables: Insights from an online retailer. *The Review of Economic Studies 82*(4), 1612–1656.

Snyder, E. M. (1956). Measuring Comparable Living Costs in Cities of Diverse Characteristics. *Monthly Labor Review*, 1187–1190.

Snyder, E. M. P. (1961). Cost of Living Indexes for Special Classes of Consumers. In *The Price Statistics of the Federal Goverment*, pp. 337–372.

Song, I. and P. K. Chintagunta (2007, November). A Discrete-Continuous Model for Multicategory Purchase Behavior of Households. *Journal of Marketing Research 54*, 595–612.

Verhoogen, E. (2008). Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector. *Quarterly Journal of Economics 123*(2), 489–530.

Waldfogel, J. (2003). Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets. *The RAND Journal of Economics 34*(3), 557–568.

Winkelried, D. and R. J. Smith (2011). Principal components instrumental variable estimation.