Are Poor Cities Cheap for Everyone?

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

Jessie Handbury

July 8, 2013

Abstract

Standard cost-of-living indexes assume that preferences are homothetic, ignoring the well-established fact that tastes vary with income. This paper considers how assuming homotheticity biases our estimates of spatial price indexes for consumers at different income levels. I use Nielsen household-level purchase data in over 500 categories of food products to calculate micro-founded income- and city-specific price indexes that account for non-homotheticity, as well as city-specific price indexes that do not. I find that the income-specific cross-city price indexes vary widely across income groups. Grocery costs are 20 percent lower in a poor city relative to a wealthy city for a low-income household, but they are 20 percent higher in the poor city for a high-income household. The homothetic price indexes perform well in predicting the cross-city variation in prices for low- and middle-income households, but poorly for high-income households.

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

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. Unfortunately, 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 across consumers.\footnote{Albouy (2009) and Moretti (2013), for example, use the ACCRA indexes to calculate real tax burdens and income inequality accounting for intra-national price variation, while Deaton (2010) and Almas (2012) use homothetic indexes based on the Penn World Table data to calculate poverty thresholds and real income inequality adjusting for international price variation.} Standard price indexes imply that a rich New Yorker would be indifferent between the set of goods available in Manhattan stores and the set available in a city will less than half the per capita income, like New Orleans. In reality, expenditure shares vary systematically with income indicating that preferences are non-homothetic.\footnote{Notable exceptions include Deaton and Dupriez (2011) who calculate country-specific poverty thresholds based on purchasing power parity deflators that reflect the consumption patterns of the global poor, and Li (2012) who uses income-specific price indexes to measure the difference in the potential welfare gains from variety for high- relative to low-expenditure households moving from rural to urban areas in India.} This paper is the first to study the implications of these non-homotheticities for regional price indexes. The results suggest that non-homotheticities in demand play a powerful role in determining the perceived price levels for rich and poor consumers. The rich find poor cities expensive because they cannot easily consume the goods they like, while the poor find high-income cities expensive because their preferred varieties are relatively expensive there. These results help to explain the divergence between high- and low-income households in their locational choices.

My initial contribution is methodological. 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 build on these micro models to develop a utility framework that represents non-homothetic preferences across many sectors of differentiated products. My model provides theoretical underpinnings to income-specific price indexes that summarize how households at different income levels value the prices and products available to them in different geographic markets. I estimate the model with data containing the purchases and demographics of approximately 40,000 U.S. households and use these estimates to calculate non-homothetic price indexes that aggregate costs across over 500 categories food products. These estimates enable me to characterize how and why the price level varies across cities in the U.S. differently for consumers at different income levels. This analysis yields three novel results.

First, I find large differences between how high- and low-income households perceive the prices and variety available in different U.S. cities. Once you account for these different preferences, markets that are relatively expensive for poor households are instead relatively cheap for the wealthy. A low-income household earning $15,000 a year faces approximately 20 percent higher grocery costs in cities with relatively high per capita income like New York relative to cities with half that per capita income like New Orleans. But the exact opposite is true for
high-income households earning $100,000 a year whose grocery costs are 20 percent lower in
the city with the higher per capita income. Standard price indexes fail to capture this systematic
variation in the living costs faced by consumers in different income classes in wealthy, relative
to poor, cities.\(^4\)

I then show that these differences are driven by cross-city variation in product variety, rather
than prices. High-income households are better off in wealthier cities because more varieties
of the high-quality products that high-income consumers prefer to consume are available in
these locations. 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.\(^5\)
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.

Finally, I demonstrate the extent to which assuming homotheticity biases our estimates of
spatial price indexes for consumers at different income levels. I find that a homothetic price
index does a better job of predicting the distribution of costs across locations for low- and
middle-income households than it does for high-income households. The homothetic price
index is highly correlated with the grocery costs for households with incomes below $70,000,
but negatively correlated with the grocery costs for households with incomes above $100,000,
systematically underestimating the costs faced by high-income consumers in wealthy relative
to poor cities. Assuming homotheticity yields much larger biases in standard cost-of-living
estimates for rich consumers than for poor consumers.

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. The starting point for these price indexes is the log-logit/constant
elasticity of substitution (CES) family of utility functions. Log-logit sub-utility functions gov-
ern how consumers allocate expenditures between products within product categories, while a
CES superstructure 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 ag-
ggregate demands for goods as if they had been derived from a representative (non-homothetic)

\(^4\)Complementary work finds large variation in inflation across income groups. Broda and Romalis (2009) use
the same Nielsen dataset, but a different methodology, to calculate income-specific U.S. inflation indexes and show
that half of the increase in conventional measures of U.S. income inequality was due to a bias caused by ignoring
the variation in consumption behavior across income groups.

\(^5\)Handbury and Weinstein (2011) 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.
This provides a way of bridging 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 amounts that consumers care about both the price and the quality of the product they are purchasing depends on their expenditure on a composite of outside, non-food products, which I assume to be normal. The intuition here is that, if high-income households spend more on cars, schooling, and mortgages, for example, then they have a greater willingness to pay for grocery products generally and, in addition, spend more on those 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) and Simonovska (2010), for example, are based on the idea that demand elasticities vary systematically with income, while Fajgelbaum et al. (2011), and Faber (2012) model non-homotheticities as a changing taste for quality.

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. food consumption. I test homotheticity against 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.

This paper also provides the first direct evidence of income-specific tastes for city-specific 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 urban consumption amenities than low-skill, low-income consumers. Previous empirical support of this theory relies on spa-

---

6 The log-logit and CES are linked mathematically such that the CES-nested log-logit utility framework yields the same aggregate outcomes as a nested-CES utility function. The origins of this result are Anderson et al. (1987), whose proof is extended to models that account for product quality in Verhoogen (2008).

7 Evidence shows that there are other reasons that demand may vary with income, related to demand for variety (Li (2012)) 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.

8 The spatial sorting of high- and low-skill workers is a salient feature of developed economies (see, e.g., Bacolod et al. (2009) and Glaeser and Resselge (2010) for the U.S., and Combes et al. (2008) for France). There is convincing evidence supporting the idea that high-skill workers co-locate in large cities because they enjoy greater productivity spillovers than low-skill workers (see, e.g., Glaeser and Mare (2001); Wheeler (2001); Davis and Din-
tial equilibrium models that assume people are perfectly mobile. Diamond (2012), for example, infers changes to skill-biased amenities as those which reconcile changes in housing price and wage data with the changing skill composition of U.S. cities between 1980 and 2000. I instead measure these amenities directly, providing cross-sectional evidence that non-housing price indexes are correlated with local incomes in such a way that encourages 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 reduced-form evidence for the mechanism behind these effects, showing that local distributors cater to local tastes. My main contribution here, however, is to demonstrate the economic significance of these externalities by measuring their impact on consumer costs.

These results have mixed implications for the question of how to account for cost-of-living differences across locations when measuring welfare. The standard use of homothetic price indexes implicitly ignores that households with different incomes have different tastes and, therefore, may perceive these relative costs differently. I find that these cost differences are large, indicating that it is 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 of real income inequality (Moretti, 2013) and the geographic distribution of real tax expenditures in the U.S. (Albouy, 2009). On the other hand, I also show that a homothetic price index represents the spatial variation in prices quite closely for households at all points but the upper tail of the income distribution.

---

9 The observed distribution of product availability is also consistent with a comparative advantage story, where skilled, high-income workers are more productive in the production of high-quality goods, independent of the fact that they like to consume these goods. While I do not differentiate between these two stories in my analysis, my demand estimation is not contaminated by any supply effect driven by comparative advantage and the resulting estimates speak to the strength of the preference externalities story.

10 Moretti (2013) shows that college graduates tend to live more in places with higher prices, so that price-adjusted welfare inequality is lower than nominal inequality. Albouy (2009) shows that we underestimate the relative tax burden of consumers with high nominal incomes when we do not account for the fact that they are disproportionately located in expensive cities. If consumers with high nominal incomes find the wealthy cities to be less expensive than the poor cities that the remainder of the population finds to be cheap, they will instead have higher real incomes and face a lower real tax burden in these locations.

11 This result is consistent with Deaton and Dupriez (2011) find that re-weighting the International Comparison Project (ICP)'s purchasing power parity (PPP) indexes to reflect the consumption patterns of the world’s poor does not change the indexes or, therefore, poverty counts dramatically.
may be reasonable to use homothetic price indexes to account for location-specific costs when calculating poverty thresholds or entitlement payments (e.g., Slesnick (2002), Deaton (2010), and Ziliak (2011)).

This paper proceeds as follows. Section 2 describes the dataset. Section 3 documents the patterns in household consumption behavior and city product availability that drive the main results of the paper. Sections 4 and 5 present the utility framework and methods that I use to model consumer behavior and measure the relative costs consumers face across cities with different prices and product availability. I estimate the parameters for four models based on this framework: one allowing for non-homotheticity in quality and price sensitivity, one allowing non-homotheticity in just one of these dimensions, and a final homothetic model. In Section 6, I present the parameter estimates for each of these models, select the model that explains the observed consumption behavior in the most parsimonious way, and analyze how the cross-city price indexes implied by the selected model vary with income. Section 7 concludes.

2 Data

The analysis in this paper is based on detailed household consumption information from the Nielsen HomeScan Panel. I use this data to infer the set of products and prices available in U.S. cities, as well as how consumers at different income levels value these products and prices. I also use MSA-level population and income data from the 2000 U.S. census to identify whether city size and wealth are correlated with prices and product availability and county-level grocery store data from the 2007 economic census in demand estimation to identify variation in prices that is related to variation in retail competition.\footnote{I describe how I build this price instrument in more detail in Section 5 below.} In what follows, I describe the structure of the Nielsen data and the key variables I draw from it.

The Nielsen HomeScan data includes information on all bar-coded food product purchases made by a demographically representative, but unbalanced, panel of over 40,000 households in 52 markets across the United States between 2003 and 2005. The households in the sample were provided with bar-code scanners and instructed to collect information such as the Universal Product Code (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, their income category, the number of household members, the ages of all members, and the occupation and education levels of the female and male head of household.

There are around 400,000 UPCs represented in the full Nielsen sample. The Nielsen data categorizes each UPC into one of 640 modules and provides detailed data on each UPC’s brand, size (including units), container, flavor, form, formula, variety, style, organic seal, and salt.
content. I determine the manufacturer of each UPC by matching the first 7 digits of the UPC code with a list of manufacturers downloaded from www.upcdatabase.com.

Product Definitions

I aggregate UPCs into a broader level of classification that I call a “product.” A product is defined as the set of UPCs within a module with the same brand, manufacturer, container size and count, diet and organic categorization, and salt content.\(^{13}\) For example, in the product module “SOFT DRINKS - CARBONATED”, there are 15 UPCs that refer to non-diet, non-organic, regular salt, and single-pack 12 ounce containers sold under the brand “COCA-COLA CLASSIC R” that are produced at “COCA-COLA USA OPERATIONS.” These 15 UPCs belong to the same product.

| Table 1: Summary Statistics for Nielsen HomeScan Data Used in Estimation |
|---|---|---|---|---|---|---|---|---|---|---|
| | Total & Per Module | Per Brand | Per Product |
| Count | Min | Median | Max | Min | Median | Max | Min | Median | Max |
| Modules | 538 | - | - | - | - | - | - | - | - |
| Brands | 12,194 | 5 | 97 | 813 | - | - | - | - | - |
| Products | 181,072 | 6 | 190 | 5,284 | 1 | 2 | 135 | - | - | - |
| UPCs | 318,825 | 6 | 289 | 9,464 | 1 | 2 | 431 | 1 | 1 | 153 |

Table 1 shows summary statistics for the sample used to estimate demand. This sample has been cleaned in three ways. Of the 640 modules included in the grocery database, 46 are for random weight items and are excluded from the analysis.\(^{14}\) Within each module, I drop any UPCs whose container size is not reported in the modal units for that module. To control for data recording errors, I drop any purchase observation for which the price paid for a UPC was greater than three times or less than a third of the median price paid per unit of any UPC within the same product or module categorization and limit the sample to products that are purchased by 20 or more households. There are between 6 and 5,284 products in each of the remaining 538 product modules. The median numbers of products and UPCs per module are 190 and 289, respectively. Although there are much fewer products in each module than there are UPCs, the modal product only contains one UPC.

The utility function presented below assumes that, conditional on price, consumers do not differentiate between UPCs in the same product. This assumption is trivial when otherwise

\(^{13}\)The container count is equal to one equal to one when each container of the product is sold individually and greater than one when multiple containers of the good are sold in a multi-pack.

\(^{14}\)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, so they are excluded from this analysis.
identical products have different UPCs because, for example, manufacturers use UPCs to track the sales of products through their respective distribution channels. The assumption is stronger in cases where different UPCs that I have defined to be the same product are differentiated by their label or flavor.

To check the extent to which consumers differentiate between UPCs within product categories, I compared the coefficient of variation for the average price paid for each UPC with the coefficient of variation for the average price paid for the set of UPCs with the same product categorization. The median UPC-level coefficient of variation is 0.14, which is only slightly lower than the median product-level coefficient of variation at 0.15. This indicates that there is little variation in the prices charged for UPCs within the same product.

**Household Income**

The Nielsen HomeScan 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 categories based on reported income. For my analysis, I exclude households with reported incomes below $10,000 and/or missing demographic data. For each of the remaining 39,767 households, I calculate a continuous, size-adjusted household income variable that represents the predicted income of each household, were it only to have one member.

To calculate size-adjusted income, I first assign each household the income 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 regress the log of this category income against dummies for household size, female and male head of household age and years of education, marital status, race, and Hispanic origin. I finally calculate each household’s log size-adjusted income by subtracting its size fixed effect from its log category income and adding back the single-member-household size fixed effect.

Figure 1 shows the distribution of the resulting size-adjusted household income across the households considered in the analysis below. The bulk of the distribution is between $10,000 and $80,000, which seems reasonable given that the per capita incomes of the cities represented in the sample range from $21,446 in New Orleans to $54,191 in San Francisco.
City-Level Product and Price Availability

I measure how consumers value the products and price available in each city in 2005 using price indexes over the set of varieties that I observe being purchased locally in that year; that is, by the households included in the sample for each market during 2005. The 2005 sample includes the purchases of 32,553 households distributed across 49 U.S. cities. Since I observe only a small sub-sample (between 116 and 1,477) of the households that reside in each city, it is likely that I observe only a sub-sample of the products available in each city. The inference that I make as to what products and prices are available in each city will be subject to biases related to the number and type of households sampled in each city.

This data limitation is common to all work that uses micro data to calculate spatial price indexes. Handbury and Weinstein (2011) show how homothetic non-parametric cross-city price indexes can be adjusted to account for potential sample size biases. Unfortunately the parametric price index methodology used in this paper does not allow for such adjustments, so I instead mitigate sample size biases using more direct sampling techniques. I discuss these in Section 5.2 below.

Nielsen groups households into 52 markets. I instead classify cities at the level of Consolidated Metropolitan Statistical Area (CMSA) where available, and the Metropolitan Statistical Area (MSA) otherwise. For example, where Nielsen classifies urban, suburban, and ex-urban New York separately, I group them all as New York-Northern New Jersey-Long Island CMSA. In the two cases in which Nielsen groups two MSAs into one market, I count the two MSAs as one city, using the sum of the population and the population-weighted per capita income. Table A.1 in Appendix E lists this data for each of the 49 cities considered in this analysis and shows how the sample households are distributed across these locations.

An alternative to household-level data for inferring the choice set in each city would be store-level data, where inference would instead be affected by biases related to the number and type of stores sampled in each city.
3 Stylized Facts

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

High-Income Households Purchase Different, More Expensive, Products than Low-Income Households  Figure 2 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 20 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 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 2 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 the difference between the price paid by high-income households for products within a module is due to the fact that they are buying different products that are sold at higher prices to all consumers.

Wealthier Cities Offer More Products that are Purchased by High-Income than by Low-Income Households  To demonstrate this point, I first calculate how much households in the highest and lowest size-adjusted income deciles spend on each product nationally. I then measure what share of high- and low-income expenditures are represented by the products in the sample for each city. Figure 3 indicates a distinct 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
Notes: 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 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 $150,000.

Notes: Numbers on plots reference the market ID of the city represented, as listed in Table A.1 in Appendix E. A similar pattern holds amongst
the sample of 23 cities (with 850 or more sample households) that are used in the price index analysis. In this sub-sample the slope is 10.3 with
a standard error of 3.1.
more heavily by low-income consumers. For example, around 0.075 more of the high-income expenditure share than the low-income expenditure share is represented in the sample for the wealthiest city, San Francisco (#45 in the figure), while around 0.015 less is represented in the sample for the poorest city, New Orleans (#24 in the figure). The systematic differences between the products available in wealthy and poor cities and, in particular, their correlation with the purchase behavior of high- and low-income households, help to explain the result that high-income households find wealthy cities to be relatively less expensive than poor cities, even though low-income households do not.

In the structural analysis below, I will formalize these stylized facts, characterizing the products that are preferred by high-income households and then quantifying how much high-income households gain from the relative abundance of these products available in wealthy, relative to poor, cities.

4 Model

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

4.1 Notation

Figure 4 shows how consumers choose to allocate expenditures. At the upper-most level, a consumer $i$ spends $W$ on a set of grocery products, denoted $G$, and $Z$ on a set of other goods, denoted $Z$, 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 non-grocery expenditures. This is generically the case as long as optimal non-grocery expenditures are normal.\(^{17}\)

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, $M = \{1, ..., M\}$, and their module expenditure $w_m$ between the varieties of grocery products in module $m$, $G_m = \{1, ..., G_m\}$, 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 $p_{mg}$

\(^{17}\)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.
(units are module-specific). I denote the set of observed grocery prices and purchase quantities for module $m$ as $\mathbb{P}_m = \{p_{mg}\}_{g \in G_m}$ and $\mathbb{Q}_m = \{q_{mg}\}_{g \in G_m}$, 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 $g \in G_m$; that is, $\sum_{g \in G_m} w_{mg} = w_m$.

### 4.2 Consumption Utility

I model consumer demand for the products in $G$ using a combination of CES and log-logit preferences. A consumer $i$’s utility from grocery consumption, conditional on their outside good expenditure $Z$, is a CES aggregate over consumer-specific module-level utilities:

$$U_{iG}(Q, Z) = \left\{ \sum_{m \in M} u_{im} (Q_m, Z)^{\frac{\sigma(Z) - 1}{\sigma(Z)}} \right\}^{\frac{\sigma(Z)}{\sigma(Z) - 1}}$$

(1)

where $\sigma(Z) > 1$ is the elasticity of substitution between modules for a consumer with outside good expenditure $Z$.

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

$$u_{im}(Q_m, Z) = \sum_{g \in G_m} u_{img}(Q_m, Z)$$  \hspace{1cm} (2)$$

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

$$u_{img}(Z) = q_{mg} \exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})$$  \hspace{1cm} (3)$$

where $\beta_{mg}$ is the quality of product $g$ in module $m$; $\varepsilon_{img}$ is the idiosyncratic utility of consumer $i$ from product $g$ in module $m$; $\gamma_m(Z)$ and $\mu_m(Z) > 0$ are weights that govern the extent to which consumers with outside good expenditure $Z$ care about product quality and their idiosyncratic utility draws.\(^{19}\)

### 4.2.1 Functional Forms

The assumption that the cross-module substitution patterns are governed by a CES utility function implies that consumers will optimally consume a positive amount in each module. In the data, the typical household buys a positive amount of a product in only 190 of 538 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 do not observe a purchase over the time period that they are in the sample.\(^{20}\) Under this scenario, households purchase a positive quantity of products in these modules in expectation. The moments used to estimate the model parameters are based on expected expenditure shares and calculated using the purchases of a market of multiple households. The fact that some households do not purchase products in certain modules during a given quarter will be reflected in the fact that these modules have low market 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.

Assuming that module utility is additive in product utilities in equation 2 implies that product-level utilities are perfectly substitutable with the utility from each of the other products within the same module. Consumers will, therefore, purchase positive quantities of only

\(^{19}\)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.

\(^{20}\)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.
the product(s) that maximize their marginal utility from product expenditure, \( \exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})/p_{mg} \). I assume that each consumer’s product-specific idiosyncratic utility draws, \( \varepsilon_{img} \), are drawn from a continuous type I extreme value distribution, with scale 0 and shape 1, so there will be a unique product that maximizes the marginal utility of expenditure for each household within each module; each household will allocate all of their module expenditure to only one product. This matches the discrete-continuous behavior of the modal household in the data, who purchases one or more units of exactly one product per module in each quarter.\(^{21}\)

### 4.2.2 Non-Homotheticities

Consumers get utility from consuming quantity \( q_{mg} \) of a product \( g \), scaled up by the exponents of the quality of the product, \( \beta_{mg} \), and their idiosyncratic utility draw for the product, \( \varepsilon_{img} \). Preferences will be non-homothetic when at least one of the weights on these scalars, \( \gamma_m(Z) \) or \( \mu_m(Z) \), or the elasticity of substitution between modules, \( \sigma(Z) \), varies with outside good expenditure and, as discussed above, this expenditure is normal. 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, \( \beta_{mg} \), for product \( g \) in module \( m \) shared by consumers who spend \( Z \) on the outside good. I assume that \( \gamma_m(Z) \) is log-linear in outside good expenditure, \( Z \), with a module specific slope, \( \gamma_m \), such that:

\[
\gamma_m(Z) = 1 + \gamma_m \ln(Z)
\]  

(4)

A consumer’s valuation for product quality in module \( m \) is increasing in \( Z \) when \( \gamma_m > 0 \).

The \( \beta_{mg} \) quality parameters are assumed to be common across products with the same brand name in a module. I employ a revealed preference approach estimating brand quality as the average willingness to pay for products with this brand across all consumers. The idea here is that product \( g \) in module \( m \) is estimated as having high quality, \( \beta_{mg} \), 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 products with the same brand as \( g \) than on products with the same brand as \( \tilde{g} \). All consumers agree on this distribution of product qualities but, for \( \gamma_m > 0 \), consumers who spend more on the outside good 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 \( \gamma_m \) would imply 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

\(^{21}\)That is, the median number of products purchased in a module by a household in a quarter is equal to one.
with a model that allows for complementarities between product quality and outside good expenditure. These complementarities imply that the elasticity of demand for quality is increasing with income, as in Hallak (2006) and Feenstra and Romalis (2012), 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, $\varepsilon_{img}$. 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:

$$\frac{1}{\mu_m(Z)} = \sigma_m(Z) - 1 \equiv \alpha_0^m + \alpha_1^m \ln(Z)$$

The inverse idiosyncratic utility draw weight is related to the elasticity of substitution between products in module $m$, $\sigma_m(Z)$, by the first equality in equation (5). For $\alpha_1^m < 0$, $\sigma_m(Z)$ 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, $\alpha_1^m < 0$ 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. I impose the same functional form assumption on the elasticity of substitution between product modules, $\sigma(Z) \equiv 1 + \alpha_0^1 + \alpha_1^1 \ln(Z)$, so, for $\alpha_1^1 < 0$, so high-$Z$ consumers will also be less sensitive to changes in the aggregate quality-adjusted price across modules.\(^{22}\)

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 $q_{\gamma}^*$ 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.

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

\(^{22}\)The role that the $\alpha$ parameters that govern the substitution elasticities play in determining the elasticity of demand with respect to price is clearly demonstrated in the expenditure share equations presented in Section 5.3 below.
explored in Aguiar and Hurst (2005) 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 2 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 food 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 (2012) or the translated additive-log utility function used in Simonovska (2010).

### 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 outside good expenditure, and their idiosyncratic utility draws. I assume that consumers draw an idiosyncratic utility \( \varepsilon_{img} \) for each product \( g \in \mathcal{G} \) 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:

\[
\sum_{m \in \mathcal{M}} \sum_{g \in \mathcal{G}_m} p_{mg} q_{mg} \leq Y_i - Z \quad \text{and} \quad q_{mg} \geq 0 \quad \forall mg \in \mathcal{G} \tag{6}
\]

The consumer’s optimal bundle in module \( m \), \( Q^*_m(w_m, Z) \), is a function of their expenditure in that module, \( w_m \), and their non-grocery expenditure, \( Z \):

\[
Q^*_m(w_m, Z) = \arg \max_{Q_m \geq 0, \text{ s.t.}} \sum_{g \in \mathcal{G}_m} q_{mg} d_{img}(Z) \tag{7}
\]

where I use \( d_{img}(Z) = \exp(\gamma_m(Z) \beta_{mg} + \mu_m(Z) \varepsilon_{img}) \) to denote the marginal utility that consumer \( i \) receives from consuming one unit of product \( g \) in module \( m \) when spending \( Z \) on non-groceries. Recall that the additive log-logit functional form implies that consumers optimally purchase a positive quantity of only one product in a module. This product maximizes their marginal utility of expenditure in a module conditional on their outside good expenditure:\(^{23}\)

\[
g^*_m(Z) = \arg \max_{g \in \mathcal{G}_m} \frac{d_{img}(Z)}{p_{mg}} \tag{8}
\]

\(^{23}\) Note that the marginal utility of expenditure in a module and, therefore, the optimal product choice, \( g^*_m \), depends on a consumer’s outside good expenditure, \( Z \), but is independent of their module expenditure, \( w_m \).
Since all of a consumer’s module expenditure, \( w_m \), is allocated to this optimal product, \( g_{im}^* \), we write the consumer’s optimal module bundle, \( Q_{im}^*(w_m, Z) \) as

\[
Q_{im}^*(w_m, Z) = (q_{i1m}^*(w_m, Z), \ldots, q_{imG_m}^*(w_m, Z))
\]

where \( q_{img}^*(w_m) = \begin{cases} 
\frac{w_m}{p_{mg}} & \text{if } g = \arg\max_{g \in G_m} \frac{d_{img}(Z)}{p_{mg}} \\
0 & \text{otherwise}
\end{cases} \) (9)

That is, a consumer \( i \) optimally consumes as much of their optimal product, \( g_{im}^*(Z) \), as their module expenditure, \( w_m \), will afford them and zero of any other product in the module. Substituting this optimal bundle into consumer \( i \)’s module \( m \) utility function, defined in equation (2), yields consumer \( i \)’s indirect utility in module \( m \) as a function of their module and outside good expenditures, \( w_m \) and \( Z \), respectively:

\[
\tilde{u}_{im}(w_m, Z) = u_{im}(Q_{im}^*(w_m, Z)) = w_m \left( \max_{g \in G_m} \frac{d_{img}(Z)}{p_{mg}} \right)
\] (10)

Consumer \( i \)’s optimal module expenditures, \( w_i^*(Z) = (w_{i1}^*(Z), \ldots, w_{iM}^*(Z)) \), can now be expressed as a function of their income, \( Y_i \), non-grocery expenditure, \( Z \), and module-level indirect utilities,

\[
w_i^*(Z) = (w_{i1}^*(Z), \ldots, w_{iM}^*(Z)) = \arg\max \left\{ \sum_{m \in M} \tilde{u}_{im}(w_m) \frac{\sigma(Z)-1}{\sigma(Z)} \right\}
\]

\[
\sum_{m \in M} w_m \leq Y_i - Z
\]

The solution to this problem, derived in Appendix B, is

\[
w_{im}^*(Z) = (Y_i - Z) \left( \frac{\max_{g \in G_m} \frac{d_{img}(Z)}{p_{mg}}}{P(P, Z, \varepsilon_i)} \right)^{\frac{\sigma(Z)}{\sigma(Z)-1}}
\]

where \( P(P, Z, \varepsilon_i) \) is a CES price index over the grocery products that a consumer \( i \) optimally consumes in each module:

\[
P(P, Z, \varepsilon_i) = \left[ \sum_{m \in M} \left( \max_{g \in G_m} \frac{d_{img}(Z)}{p_{mg}} \right)^{\frac{\sigma(Z)}{\sigma(Z)-1}} \right]^{\frac{1}{1-\sigma(Z)}}
\]

Plugging the optimal product choices and module expenditures derived above into the direct utility function defined in equations (1)-(3) yields the indirect utility of consumer \( i \) from grocery
consumption:

\[ V(\mathbb{P}, Y_i, Z, \varepsilon_i) = \frac{(Y_i - Z)}{P(\mathbb{P}, Z, \varepsilon_i)} \]  \quad (14)

5 Empirical Strategy

The main goal of this paper is to measure how variation in the prices and product availability across U.S. cities differentially impacts the utility of consumers at different income levels. 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, I describe 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 \( \mathbb{P} \). This paper seeks to compare the utility consumers get from the prices and products available to them in different markets, so I now introduce a market subscript to equation (14), writing the indirect utility of a consumer \( i \) in market \( t \) as

\[ V(\mathbb{P}_t, Y_i, Z_{it}, \varepsilon_i) = \frac{(Y_i - Z_{it})}{P(\mathbb{P}_t, Z_{it}, \varepsilon_i)} \]  \quad (15)

where the set of prices and products available to household \( i \), \( \mathbb{P}_t = \{p_{mgt}\}_{g \in G_t} \), 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, \( \varepsilon_i \). I wish to study the systematic variation in utility across consumers earning different incomes and am interested in the latter two channels only insofar as non-grocery expenditures \( Z_{it} \) and/or idiosyncratic utility draws \( \varepsilon_i \) are correlated with income.

The idiosyncratic utility \( \varepsilon_i \) draws are, by definition, uncorrelated with consumer income \( Y_i \), introducing only random variation in consumers’ indirect utilities.\(^{24}\) The most direct way to abstract from this random variation would be to take the expectation of the indirect utility defined in equation (15) over the idiosyncratic draws. Unfortunately, there is no analytic solution to this problem, and numerical solutions are computationally intensive.\(^{25}\) I instead measure the

\(^{24}\)The fact that the idiosyncratic utilities are drawn from a continuous distribution implies that no two consumers, even with the same income, \( Y_i \), will get the same utility from a set of consumption opportunities, \( \mathbb{P}_t \).

\(^{25}\)For robustness, I numerically integrated over the \( \varepsilon_i \) draws in price indexes for the highest- and lowest-income
relative utility of households at various income levels across different markets by measuring the
utility of an income-specific representative consumer.

I assume that the representative consumer’s utility from consuming a grocery bundle $Q$ is a
nested-CES function conditional on their outside good expenditure $Z$ defined as:

$$U_{CES}^G(Q, Z) = \left\{ \sum_{m \in M} \left[ \sum_{g \in G_m} \left[ q_{mg} \exp(\beta_{mg} \gamma_m(Z)) \right] \right] \frac{\sigma_m(Z)}{\sigma(Z)} \left( \frac{\sigma(Z)}{\sigma_m(Z)} - 1 \right) \right\} \frac{\sigma(Z)}{\sigma_m(Z)} - 1, \quad (16)$$

where $q_{mg}$, $\beta_{mg}$, $\gamma_m(Z)$, $\sigma_m(Z)$, and $\sigma(Z)$ take the same definitions as in the log-logit utility
function presented in Section 4 above.26 The indirect utility of this representative consumer
from income $Y_i$ and prices and products $P_t$, $V_{CES}^i(P_t, Y_i)$, takes a similar form to the indirect
utility of the idiosyncratic consumer provided in equation (15) 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:

$$V_{CES}^i(P_t, Y_i, Z_{it}) = \frac{(Y_i - Z_{it})}{P_{CES}(P_t, Z_{it})}, \quad (17)$$

where

$$P_{CES}(P_t, Z_{it}) = \left[ \sum_{m \in M} \left( \left[ \sum_{g \in G_m} \left( \exp(\beta_{mg} \gamma_m(Z_{it})) \right) \frac{1 - \sigma_m(Z_{it})}{1 - \sigma_m(Z_{it})} \right] \right)^{1 - \sigma(Z_{it})} \right]^{1 - \sigma(Z_{it})}$$

While the indirect utility of the idiosyncratic consumer was consumer-specific through three
variables ($Y_i$, $Z_{it}$, and $\varepsilon_i$), the representative consumer indirect utility defined in equation (17)
is consumer-specific through only $Y_i$ and $Z_{it}$. 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, $Y$, such that $Z_{it} = Z(Y)$. Theoretically, this assumption could be violated since con-
sumers 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 in-
markets in the full Nielsen sample: San Francisco and New Orleans. Using the numerical integration method,
I found a gap of approximately 60 percent between the relative grocery costs faced by the average high-income
consumer between these markets and those face by the average low-income consumer. Using the CES method,
I found similar results; the high-income representative consumer faces 75 percent lower grocery costs in San
Francisco relative to New Orleans relative to low-income representative consumer.

26 In Appendix C, I show that this income-specific, nested, asymmetric CES utility function yields identical
within-grocery budget shares as the CES-nested log-logit utility function that I estimate.
come is surprisingly consistent across cities. Figure 5 demonstrates that households earning higher incomes spend a smaller share of their income on grocery products but, within income groups, the average grocery expenditure share does not vary greatly across cities.

Figure 5: Income-Specific Food Expenditure Shares Across Markets

Note: For the purposes of visual clarity, only a representative sample of deciles are represented. The coefficient of variation of household grocery expenditure shares is 78 across all households in the sample, but drops to between 32 and 52 when you only consider households within each income decile.

Under this assumption, we can express the consumer’s indirect utility as a function of market prices, $P_t$, and consumer income, $Y_i$ alone:

$$V^{CES}(P_t, Y_i) = \frac{(Y_i - Z(Y_i))}{P^{CES}(P_t, Z(Y_i))},$$

where

$$P^{CES}(P_t, Z(Y_i)) = \left[ \sum_{m \in M} \left( \sum_{g \in G_{mt}} \left( \exp(\beta_{mg} \gamma_m(Z(Y_i))) \right) \frac{1 - \sigma_m(Z(Y_i))}{1 - \sigma_m(Z(Y_i))} \right) \right]^\frac{1}{1 - \sigma(Z(Y_i))}$$

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:

$$\frac{V(P_t, Y_i)}{V(P_{t'}, Y_i)} = \frac{P^{CES}(P_{t'}, Z(Y_i))}{P^{CES}(P_t, Z(Y_i))}$$

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 CES price index defined in equation (19) is, therefore, central to the main results of this paper and the remainder of this section will explain how I measure this price index for households at different income levels in different cities across the U.S..

Before proceeding, it is worth taking a moment here to note that this approach to measuring income-specific spatial price indexes is different from the approach that Broda and Romalis (2009) use to calculate income-specific inflation with the same Nielsen HomeScan data. Broda and Romalis (2009) 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, $\sigma_m$ and $\sigma$, in the framework above, to vary with income. It is, however, more flexible in that Broda and Romalis (2009) implicitly allow for households at different income levels to have entirely different revealed preferences ($\beta_{mg}$) 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 (2012), 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).\footnote{The Feenstra and Romalis (2012) 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.}

5.2 Inferring Prices and Product Availability

The first input to the price index defined in equation (19) is a market-specific price vector, $P_t$, representing the set of prices and products available to consumers in a city $t$. I proxy for the set of prices and products available to consumers in each city in 2005 using the set of products and unit values represented in the purchases of the households surveyed by Nielsen in that year. The 2005 HomeScan dataset includes the purchases of between 100 and 1,500 households in each of the 49 MSAs included in the analysis. I deal with potential biases that could result from correlations between the number and type of households in the Nielsen sample for each city in two ways. Each method relies on the random selection sub-samples of households whose purchases I consider when constructing the $P_t$ price vectors for each city.

The first potential concern here is that the number of households I observe varies systematically across cities with different per capita incomes. The Nielsen HomeScan sample tends to include more households in larger, wealthier cities, such that I might observe more product variety available and mechanically estimate lower relative price levels in these locations. The
correlations of the sample household count with market population and market income are 0.57 and 0.43, respectively. I mitigate sample size biases in the price indexes I calculate by limiting the number of households whose purchases I consider to be equal across cities. In my main analysis, I build the price vector using the purchases of 850 randomly-selected households in each of the 23 cities with 850 or more households in the Nielsen sample.

An additional concern is that Nielsen samples a demographically-representative set of households in each city, so I observe the purchases of more high-income households in the samples for high-income cities than for low-income cities. Since tastes are identified using the same sample of purchases, the sampling pattern might lead me to conclude that wealthy cities have more varieties that favor wealthy tastes, simply because I have identified these tastes using the purchases of the wealthy households who are disproportionately sampled in wealthy cities. I deal with this potential sample bias in a robustness check where I infer each city’s price and product availability using only the purchases of a stratified sample of households in each city. These stratified samples include a randomly-selected set of 570 households, 190 from each tercile of the full-sample income distribution, for each of the 22 cities for which I observe 190 or more households in each tercile (i.e., every city in which I observe 850 or more households in total except for San Francisco, where I only observe 145 low-income households).

5.3 Parameter Estimation

The second set of inputs into the price index defined in equation (19) 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 $\theta$ defined as

$$\theta = \{(\beta_1, \ldots, \beta_M), (\gamma_1, \ldots, \gamma_M), (\alpha_1, \ldots, \alpha_M), \alpha^0, \alpha^1\}$$

where $\alpha_m = \{\alpha^0_m, \alpha^1_m\}$ and $\beta_m = \{\beta_{m1}, \ldots, \beta_{mG_m}\}$.

To calculate the price index, I estimate these parameters in a two-stage GMM procedure minimizing the distance between expenditure shares observed for market-specific groups of households with similar income levels and the model’s predictions for these expenditure shares. The moments used in this analysis are based on two types of relative expenditure share equations. The first defines the market share of a product within a module and the second defines the market share of a module in total grocery expenditures.

5.3.1 Within-Module Estimating Equation

The within-module expenditure share on product $g$ in module $m$ for a group of households with the same outside good expenditure, $Z_i$, facing a common vector of module prices, $\mathbb{P}_m$, derived
in the Appendix B, is
\[
s_{mg|m}(Z_i, \mathbb{P}_m) = \mathbb{E}_e[s_{mg|m}] = \frac{\exp[(\alpha^0_m + \alpha^1_m \ln(Z_i))(\beta_{mg}(1 + \gamma_m \ln(Z_i)) - \ln p_{mg})]}{\sum_{g' \in G_m} \exp[(\alpha^0_m + \alpha^1_m \ln(Z_i))(\beta_{mg'}(1 + \gamma_m \ln(Z_i)) - \ln p_{mg'})]}
\]

The denominator of this market share will not vary across products within a module \(m\) and, therefore, drops out when I take the log of this expenditure share for any product \(g\) in module \(m\) and difference from the log expenditure share for a fixed product \(\bar{g}_m\) in the same module:
\[
\ln \left( \frac{s_{mg|m}(Z_i, \mathbb{P}_m)}{s_{m\bar{g}_m|m}(Z_i, \mathbb{P}_m)} \right) = (\alpha^0_m + \alpha^1_m \ln(Z_i)) \left[ (\beta_{mg} - \beta_{m\bar{g}_m})(1 + \gamma_m \ln(Z_i)) - (\ln p_{mg} - \ln p_{m\bar{g}_m}) \right]
\]
\[(21)\]

Equation (21) defines the expected within-module expenditure share of a set of households with outside good expenditure \(Z_i\) facing prices \(p_{mg}\) and \(p_{m\bar{g}_m}\) on product \(g\) relative to product \(\bar{g}_m\) in terms of module-specific price sensitivity parameters \((\alpha^0_m, \alpha^1_m)\) and quality taste-income gradients \((\gamma_m)\), as well as, product-specific relative quality parameters \((\beta_{mg} - \beta_{m\bar{g}_m})\) for each product \(g \in G\). I denote this set of parameters by \(\theta_1 = \{\theta_{1m}\}_{m \in \mathbb{M}}\) where
\[
\theta_{1m} = \left\{ \alpha^0_m, \alpha^1_m, \gamma_m, \{\beta_{mg} - \beta_{m\bar{g}_m}\}_{g \in G_m} \right\}
\]
for each module \(m \in \mathbb{M}\).

To estimate these parameters, I proxy outside good expenditure, \(Z_i\), with household income, \(Y_i\) and split the sample of households for each quarter-MSA market, \(t\), into income quintiles, denoted by \(k\). I then construct moments based on the following sample analog to equation 21:
\[
\ln \left( \frac{s_{kmgt}}{s_{k\bar{g}_m|t}} \right) = (\alpha^0_m + \alpha^1_m y_{kt}) \left[ (\beta_{mg} - \beta_{m\bar{g}_m})(1 + \gamma_m y_{kt}) - \ln \left( \frac{p_{kmgt}}{p_{k\bar{g}_m|t}} \right) \right] + \nu_{k\bar{g}_m|t}
\]
\[(22)\]

where \(s_{kmgt}\) and \(s_{k\bar{g}_m|t}\) are the observed expenditure share of income quintile \(k\) in market \(t\) on products \(g\) and \(\bar{g}_m\) in module \(m\); \(p_{kmgt}\) and \(p_{k\bar{g}_m|t}\) are the unit values paid by households in income quintile \(k\) in market \(t\) for products \(g\) and \(\bar{g}_m\); and \(y_{kt}\) is the median log income of households in each income quintile \(k\) in market \(t\). \(\bar{g}_m \in \mathbb{G}_m\) is a fixed base product for

---

28 The utility function assumes weak separability between modules and the independence of irrelevant alternatives (IIA) property both across modules and across products with the same quality parameter. Although neither of these are realistic characteristics of consumer behavior, they are useful for the purposes of estimation as they imply that relative market expenditure shares can be derived as functions of observed variables, such as household income, expenditures, and transaction prices.

29 Recall that I restrict \(\beta_{mg}\) to be equal across all products \(g\) with the same brand name.

30 To ensure that the purchases included in expenditure shares are representative of households’ expenditures over the entire quarter, I exclude households in any quarter if they do not report over a period of over two weeks in the quarter.

31 In a slight abuse of notation, I denote the coefficients on log income using the same notation used for the
module $m$ which I define to be the product that is sold in the largest number of markets in its module. $\nu_{kgm,t}$ is the error in the predicted value of the relative product shares. This error includes differences between the mean prices paid and median incomes of households in quintile $k$ in market $t$ ($p_{kmqt}$ and $y_k$) and the actual prices observed and incomes earned by these households. This error also includes measurement error in the collection of the raw data.

The moments are based on the assumption that $\mathbb{E}[Z_1 \cdot \nu] = 0$ for a set of instruments $Z_1$. These instruments include a set of brand dummies, a set of price instruments, and interaction terms between these sets of variables and market-quintile income, $y_{kt}$. The set of brand dummies includes one dummy for each brand except the base product in each module $\bar{g}_m$. I do not use prices as instruments because they might be correlated with the error term, $\nu_{kgm,t}$, across markets through market-specific product tastes, or across products within brands through product-specific tastes, neither of which are accounted for in the model. Instead I use a range of measures that capture temporal cost shocks that are correlated either across geography or over time as well as spatial shocks in the level of market competition.

The main price instruments are similar to those used in Hausman et al. (1994) and Nevo (2000). First, I use the average price paid by consumers in the same income quintile in all other geographic markets in the same time period. I expect that this instrument will be correlated with price through national cost shocks, such as the increase in the price of wheat. Second, I use the average price paid by consumers in the same income quintile in all geographic markets in the same region in the same time period. This proxy captures any regional cost shocks and relies on the assumption that market-specific taste shocks are not shared across regional MSAs.

Additionally, I use the price paid for a product by consumers in the same income quintile in the same market in both the lead and lag time quarter. These two instruments are similar to those used in Asker (2005) and are intended to capture persistent local cost shocks, such as increases in sales taxes or wages. The strength of this instrument relies on MSA-specific cost shocks that persist over more than one quarter and its validity relies on the assumption that MSA-specific demand shocks do not persist for more than one quarter. I also use county-level data on the number of grocery stores per capita and the number of products sold in a market

---

32 Note that, since the estimation is based on a relative share equation, I do not need to observe every product available to these households in order to calculate the moment condition. I do require, however, that the base product, $\bar{g}_m$, is purchased by a household in a given income decile-market in order for that decile-market’s purchases to be included in the within-module estimation procedure for module $m$.

33 I deal with problems associated with using a large number of potentially weak instruments by constructing $Z_1$ using principal component analysis, as suggested in Bai and Ng (2010). The $Z_1$ used in estimation is the set of factors that explain the 99.5 percent of the variation in the full set of price instruments, brand dummies, and the interactions of both with quintile income.

34 The region categorizations are provided in Table A.1 in Appendix E.
within the relevant product module to capture any price variation related to the level of retail and product competition in each module.\textsuperscript{35}

The $\theta_1$ parameters are estimated in separate non-linear GMM procedures for each module with optimal weighting matrices calculated in the conventional two-step procedure. Relative brand quality, $\beta_{mg} - \beta_{m\bar{g}m}$, is identified by variation in the average market shares of products within a brand, relative to the market share of the brand of the base product $\bar{g}_m$, conditional on price. The idea here is that, if products with two different brands sell at the same price, but products under brand A have higher average relative market shares across all income quintiles and MSA-quarter markets than products under brand B, then brand A will be assigned a higher quality parameter relative to the brand of the base product in the module. The quality-income gradient $\gamma_m$ parameters that govern how demand for quality varies with price are identified by the extent to which the relative market shares of high-income quintile households are even more biased towards products under “high-quality” brands, i.e., those that have higher market shares across all income quintiles, than low-income quintile market shares. Conditional on brand quality, the base price sensitivity $\alpha^0_m$ parameter is identified by the extent to which relative market shares co-vary with the components of relative price variation captured by the price instruments, and the $\alpha^1_m$ parameter is identified by the extent to which the relative market shares of high-income quintiles co-vary more (or less) with these relative price changes.

To test the strength of the price instruments, I compare the price elasticities estimated for the lower-level of the demand system using different subsets of the instruments described above, as well as using non-linear least squares (NLLS), ignoring the potential price endogeneity. This analysis yields two main results. First, the price coefficients are sensitive to the use of price instruments: ignoring price endogeneity in the NLLS specification, attenuates the base, $\alpha^0_m$, price coefficient for the typical module by between 25 and 26 percent, depending on the parameter restrictions imposed. Second, the price coefficients are primarily identified by variation in the first two national and regional cost shock instruments. Excluding all of the remaining instruments changes the typical base, $\alpha^0_m$, price coefficient estimates by between 0.75 and 0.85 percent, depending on the instruments excluded and parameter restrictions imposed.

5.3.2 Across-Module Estimating Equation

The remaining model parameters, $\alpha^0$, $\alpha^1$, and $\{\beta_{g_m}\}_{g \in G_m}$, are identified using moments based on the model’s prediction for module-level market shares. Specifically, the expected log expenditure share in module $m$ relative to $\bar{m}$ for a group of households with the same outside good

\textsuperscript{35}In each quarter-MSA market $t$, consumers in a given income quintile will purchase products in stores in a range of counties within their MSA. To construct the retail competition instrument, I take the weighted average of the county-level per capita store count in an MSA, where weights are determined by the purchases that are observed of a product by consumers in the relevant income quintile in the relevant MSA-quarter market.
expenditure, $Z_i$, facing a common vector of grocery prices, $\mathbb{P}$, is derived in Appendix B to be equal to

$$E_c [\ln s_{im} - \ln s_{i\bar{m}}] = - (\alpha^0 + \alpha^1 \ln(Z_i)) [\ln V_m(Z_i, \mathbb{P}_m) - \ln V_{\bar{m}}(Z_i, \mathbb{P}_{\bar{m}})]$$  \hspace{1cm} (23)

where $V_m(Z_i, \mathbb{P}_m)$ is a CES-style index over price-adjusted product qualities:

$$V_m(Z_i, \mathbb{P}_m) = \left[ \sum_{g \in G_m} \left( \exp(\beta_{mg}(1 + \gamma_{m} \ln Z_i)) \right) \right]$$

$$\frac{1 - (\alpha^0_m + \alpha^1_m \ln(Z_i))}{p_{mg}}$$

$$- (\alpha^0_m + \alpha^1_m \ln(Z_i))$$

\hspace{1cm} (24)

Together equations (23) and (24) define the expected relative module expenditure share of a set of households with outside good expenditure $Z_i$ that face prices $\mathbb{P}_m$ and $\mathbb{P}_{\bar{m}}$ in terms of parameters $\alpha^0$ and $\alpha^1$, as well as $\alpha_m$, $\gamma_m$, $\beta_{mg}$ for all $g \in G_m$, and $\alpha_{\bar{m}}$, $\gamma_{\bar{m}}$, $\beta_{\bar{mg}}$ for all $g \in G_{\bar{m}}$.

I use these expressions as the basis for calculating moments for each module $m \neq \bar{m}$. These moment are, in turn, used to estimate the cross-module substitution parameters, $\alpha^0$ and $\alpha^1$, as well as the quality of the base product in each module, $\beta_{m\bar{m}}$, for all modules $m \in \mathbb{M}$, except for the base module $\bar{m}$.\(^{36}\) I denote this set of parameters by $\theta_2$:

$$\theta_2 = \left\{ \alpha^0, \alpha^1, \left\{ \beta_{m\bar{m}} \right\}_{m \in \mathbb{M}, m \neq \bar{m}} \right\}$$

To estimate these parameters, I again proxy for outside good expenditures with household income and aggregate the household-level purchase data to the income-quintile, MSA-quarter level and relate the difference between quintile $k$’s expenditure on module $m$ expenditure relative to module $\bar{m}$ in market $t$ to the model prediction for this relative share in the following sample analog to equation (23):

$$\ln \left( \frac{s_{kmt}}{s_{k\bar{m}t}} \right) = (\alpha^0 + \alpha^1 y_{kt}) \left[ \ln V_m(y_{kt}, \mathbb{P}_{kmt}) - \ln V_{\bar{m}}(y_{kt}, \mathbb{P}_{k\bar{m}t}) \right] + u_{km\bar{m}t}$$

\hspace{1cm} (25)

where $s_{kmt}$ and $s_{k\bar{m}t}$ are grocery expenditure shares that households in income quintile $k$ in market $t$ spend on modules $m$ and $\bar{m}$, respectively, and $V_m(y_{kt}, \mathbb{P}_{kmt})$ and $V_{\bar{m}}(y_{kt}, \mathbb{P}_{k\bar{m}t})$ are the price-adjusted quality indexes, or inclusive values, that summarize the indirect utility that income quintile $k$ households get from each unit of expenditure on modules $m$ and $\bar{m}$, respectively. The inclusive value a quintile $k$ faces for module $m$ in market $t$ is defined as a function of the quintile-market income ($y_{kt}$), the vector of module-$m$ prices and products recorded in the purchases of households in this quintile $k$ in market $t$ ($\mathbb{P}_{kmt}$), and module-level utility parame-

\(^{36}\)I normalize the quality of the base product in the base module to equal zero.
The estimation procedure described in Section 5.3.1 above yields estimates for all of the module-level parameters except for the base-product qualities. I can re-write estimating equation (25) to reflect that the second set of sample moments will be constructed using these first-stage estimates, which I will denote \( \hat{\theta}_1 \), as follows:

\[
\ln \left( \frac{s_{\text{kmt}}}{s_{\text{k\tilde{m}t}}} \right) = (\alpha^0 + \alpha^1 y_{kt}) \left[ \Delta V_{1m\tilde{m}}(y_{kt}, \bar{P}_{\text{kmt}}, \bar{P}_{\text{km\tilde{t}}}, \hat{\theta}_1) + \beta_{m\tilde{m}n}(1 + \gamma_m y_{kt}) \right] + \nu_{\text{km\tilde{m}t}} + u_{\text{kmt}} \quad (27)
\]

where

\[
\Delta V_{1m\tilde{m}}(y_{kt}, \bar{P}_{\text{kmt}}, \bar{P}_{\text{km\tilde{t}}}, \theta_1) = \ln V_m(y_{kt}, \bar{P}_{\text{kmt}}) - \ln V_m(y_{kt}, \bar{P}_{\text{km\tilde{t}}}) - \beta_{m\tilde{m}n}(1 + \gamma_m y_{kt})
\]

is the difference between the log inclusive values for modules \( m \) and \( \tilde{m} \), adjusted for the perceived quality of the module \( m \) base product, \( \tilde{g}_m \). The \( u_{\text{km\tilde{m}t}} \) errors refer to the difference between the observed and predicted values of relative module shares in each market when the predicted values are calculated using the true values for the \( \theta_1 \) parameters. I use \( \nu_{\text{km\tilde{m}t}} \) to denote the additional errors in the sample predicted values due to the fact that the first-stage parameters are estimated with error such that \( \Delta V_{1m\tilde{m}}(y_{kt}, \bar{P}_{\text{kmt}}, \bar{P}_{\text{km\tilde{t}}}, \theta_1) \neq \Delta V_{1m\tilde{m}}(y_{kt}, \bar{P}_{\text{kmt}}, \bar{P}_{\text{km\tilde{t}}}, \hat{\theta}_1) \) and \( \gamma_m \neq \tilde{\gamma}_m \).

I estimate the \( \theta_2 \) parameters using a single two-step non-linear GMM procedure based on the assumption that \( \mathbb{E} [Z_2 \cdot (\nu + u)] = 0 \) for a set of instruments \( Z_2 \). These instruments include a set of dummies for all modules except the base module, two instruments for \( \Delta V_{1m\tilde{m}}(y_{kt}, \bar{P}_{\text{kmt}}, \bar{P}_{\text{km\tilde{t}}}, \hat{\theta}_1) \) calculated using the same national and regional cost shock instruments for prices that are used in the module-level estimation, and each of these dummy and cost-shock instruments interacted with income, \( y_{kt} \).

The base \( \alpha^0 \) substitution elasticity parameter is identified by the extent to which relative module shares react to national and regional module-level cost shocks. The \( \alpha^1 \) parameter gov-

---

37 This expression is derived in Appendix B.

38 The national and regional cost shock instruments are differenced quality-adjusted price indexes are constructed in the same way as \( \Delta V_{1m\tilde{m}}(y_{kt}, \bar{P}_{\text{kmt}}, \bar{P}_{\text{km\tilde{t}}}, \hat{\theta}_1) \), but using different price vectors: \( \Delta V_{1m\tilde{m}}(y_{kt}, \bar{P}_{\text{kmt}}, \bar{P}_{\text{km\tilde{t}}}, \hat{\theta}_1) \) and \( \Delta V_{1m\tilde{m}}(y_{kt}, \bar{P}_{\text{kmt}}, \bar{P}_{\text{km\tilde{t}}}, \hat{\theta}_1) \), respectively. In place of the mean prices paid by income quintile \( k \) households in geographic-temporal market \( t \), \( \bar{P}_{\text{km\tilde{t}}} \) and \( \bar{P}_{\text{km\tilde{t}}} \) include mean prices paid by income quintile \( k \) households in all other geographic markets the same time period for the same set of module \( m \) and \( \tilde{m} \) products purchased by income quintile \( k \) households in market \( t \), \( \bar{P}_{\text{km\tilde{t}}} \) and \( \bar{P}_{\text{km\tilde{t}}} \) instead replace the mean prices paid by income quintile \( k \) households in market \( t \) with the mean prices paid for the same set of products by income quintile \( k \) households in all other cities in the same region in the same time period.
erning how much more or less substitutable high-income consumers find products in different modules is identified by the extent to which high- relative to low-income quintile module shares are more or less reactive to these shocks.

The relative inclusive value, $\Delta V_{1m\bar{m}}$, function is scaled up or down by the quality of the brand of the base product, $\bar{g}_m$, in a module $m$ relative to the quality of the brand of the base product, $\bar{g}_{\bar{m}}$, in the base module $\bar{m}$, milk, which is normalized to equal zero. Any difference between the relative expenditure share of module $m$ relative to milk and what would be expected given the relative inclusive value of the two modules and the estimates of the $\alpha^0$ and $\alpha^1$ parameters will identify the quality of the brand of the base product in the module, $\beta_{m\bar{g}_m}$. Together with the relative brand quality estimates from the first stage of estimation, $\beta_{mg} - \beta_{m\bar{g}_m}$, the base brand quality estimates define the quality of each brand in the dataset relative to the quality of the base brand in the milk module.

This estimation procedure yields consistent estimates for $\theta_2$, but the variance-covariance matrix of these parameters will be biased due to the presence of the first-step estimates for $\theta_1$ in the $\nu$ component of the error. I adjust this variance-covariance matrix to account for the errors from the first stage of the estimation following the GMM analog of the Murphy and Topel (1985) procedure outlined in Newey and McFadden (1994). The adjusted variance-covariance matrix yields consistent standard errors for the $\theta_2$ estimates.

### 6 Results

#### 6.1 Parameter Estimates

The model was estimated 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 2 summarizes the estimates for the module-level parameters in each of these four models the 504 modules for which estimates were obtained. Panel B summarizes the subset of parameter estimates that are statistically significant at the 95 percent level.

Columns [1] through [3] of each table summarize the parameter estimates for the unrestricted version of the model. The corresponding utility function that governs consumer $i$’s preferences between products $g \in \mathbb{G}_m$ within each module $m$ is obtained by subbing the parametrizations for $\gamma_{im}(Z)$ and $\mu_{im}(Z)$, provided in equations (4) and (5), respectively, into equation (2):

---

39 Appendix D details how these adjustments are calculated.

40 Statistical significance implies that the lower bound of the 95 percent confidence interval for the $\sigma_{im}$ estimates is greater than one or that the lower (upper) bound of the 95 percent confidence intervals for $\alpha_{im}$ and $\gamma_{im}$ estimates are greater (less than) zero.
\[ u_{im}(Q_m, Z) = \sum_{g \in G_m} q_{mg} \exp \left( \beta_{mg}(1 + \gamma_m \ln Z) + \frac{\varepsilon_{img}}{\alpha_0^m + \alpha_1^m \ln Z} \right), \]  

where \( \beta_{mg} \) characterizes the quality of product \( g \) relative to the base product in the milk module and \( \varepsilon_{img} \) is consumer \( i \)'s idiosyncratic utility from product \( g \). In this model, the weights that consumers place on quality and their idiosyncratic utility when determining their product ranking are functions of their non-grocery expenditure, \( Z \), the log of which is proxied in estimation by log consumer income, \( y \). The estimate for the elasticity of substitution between products is a function of the estimated weight placed on the idiosyncratic utility, \( \hat{\sigma}_m(y) = 1 + \hat{\alpha}_0^m + \hat{\alpha}_1^m y \). Therefore, when the estimated value for this weight varies with income, or \( \alpha_1^m \neq 0 \), the elasticity of substitution will also vary with income. The first column of Table 2 reports the elasticity of substitution of a consumer with the mean log income level in the sample for each module, or \( \hat{\sigma}_m = 1 + \hat{\alpha}_0^m + \hat{\alpha}_1^m \bar{y}_m \). The median of this elasticity is 2.09, with an inter-quartile range of 1.58 to 2.59. The magnitude and distribution of these estimates is similar across all four models. These estimates imply a median price elasticity of -1.09 and an inter-quartile range of -0.58 to -1.59. These parameter estimates are well-identified in most modules, with at least 380 out of 504 significant at the 95 percent level in all four models.

### Table 2: Summary Statistics for Parameter Estimates

#### Panel A: All Estimates

<table>
<thead>
<tr>
<th>Model: NH in Quality and Price</th>
<th>NH in Quality</th>
<th>NH in Price</th>
<th>Homothetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restrictions: None</td>
<td>( \alpha_1^m = 0 )</td>
<td>( \gamma_m = 0 )</td>
<td>( \gamma_m = 0 ) &amp; ( \alpha_1^m = 0 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter:</th>
<th>( \bar{\sigma}_m )</th>
<th>( \gamma_m )</th>
<th>( \alpha_1^m )</th>
<th>( \bar{\sigma}_m )</th>
<th>( \gamma_m )</th>
<th>( \alpha_1^m )</th>
<th>( \bar{\sigma}_m )</th>
<th>( \gamma_m )</th>
<th>( \alpha_1^m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>504</td>
<td>504</td>
<td>504</td>
<td>504</td>
<td>504</td>
<td>504</td>
<td>504</td>
<td>504</td>
<td>504</td>
</tr>
<tr>
<td>p25</td>
<td>1.58</td>
<td>-0.04</td>
<td>-0.19</td>
<td>1.62</td>
<td>-0.08</td>
<td>1.66</td>
<td>-0.03</td>
<td>1.64</td>
<td>1.64</td>
</tr>
<tr>
<td>p50</td>
<td>2.09</td>
<td>0.15</td>
<td>-0.05</td>
<td>2.13</td>
<td>0.09</td>
<td>2.14</td>
<td>0.04</td>
<td>2.11</td>
<td>2.11</td>
</tr>
<tr>
<td>p75</td>
<td>2.59</td>
<td>0.30</td>
<td>0.05</td>
<td>2.60</td>
<td>0.19</td>
<td>2.62</td>
<td>0.15</td>
<td>2.61</td>
<td>2.61</td>
</tr>
</tbody>
</table>

#### Panel B: Statistically Significant Estimates

<table>
<thead>
<tr>
<th>Model: NH in Quality and Price</th>
<th>NH in Quality</th>
<th>NH in Price</th>
<th>Homothetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restrictions: None</td>
<td>( \alpha_1^m = 0 )</td>
<td>( \gamma_m = 0 )</td>
<td>( \gamma_m = 0 ) &amp; ( \alpha_1^m = 0 )</td>
</tr>
</tbody>
</table>

| Parameter: | \( \bar{\sigma}_m \) | \( \gamma_m \) | \( \alpha_1^m \) | \( \bar{\sigma}_m \) | \( \gamma_m \) | \( \alpha_1^m \) | \( \bar{\sigma}_m \) | \( \gamma_m \) | \( \alpha_1^m \) | \( \bar{\sigma}_m \) |
|------------|------------------|---------------|---------------|------------------|---------------|---------------|------------------|---------------|---------------|------------------|---------------|
| Count      | 383              | 254           | 171           | 380              | 301           | 393           | 263              | 391           | 391           |
| p25        | 1.87             | 0.15          | -0.31         | 1.93             | 0.05          | 1.88          | -0.05            | 1.88          | 1.88          |
| p50        | 2.27             | 0.23          | -0.19         | 2.28             | 0.15          | 2.26          | 0.10             | 2.26          | 2.26          |
| p75        | 2.72             | 0.35          | -0.10         | 2.73             | 0.23          | 2.73          | 0.19             | 2.69          | 2.69          |
We can get some sense of the reasonableness of these point estimates by comparing them to those found in other studies. The own-price elasticities found here are smaller than those estimated in Nevo (2000), who finds the own-price elasticity across cereal products to be between -2.2 and -4.2 whereas I estimate the own-price elasticity for cereals to be -1.13. On the other hand, the own-price elasticities here are larger than those estimated in Dube (2004), who finds the own-price elasticity of demand for carbonated beverages to be between -0.42 and -0.85, whereas I find the own-price elasticity for carbonated beverages to be -1.96.\footnote{All of these estimates are much lower in magnitude than the own-price elasticities implied by the elasticity of substitution estimates in Broda and Weinstein (2008), who use the Feenstra (1994) methodology to identify the elasticity of substitution between products. When I estimate the above model with this method, the median, as implied by the elasticity of substitution estimates for the largest 100 modules by sales value, rises from 2.4 to 7.7. I do not use this method more broadly because it does not allow me to identify the brand quality \( \beta_{mg} \) parameters. The Feenstra (1994) method also relies on the restrictive assumption that all products in the same module share the same price elasticity of supply to help identify the price elasticity of demand.}

Columns [2] and [5] of Table 2 summarize the distribution of the estimated values for \( \gamma_m \). All four models assume that all consumers agree on the relative quality of products, as described by the distribution of the \( \beta_{mg} \) parameters for products \( g \in G_m \) within a module \( m \). For positive values of \( \gamma_m \), 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 outside good expenditure \( Z \). This implies that consumers with higher expenditures on the outside good 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 \( \beta_{mg} \) estimates, that is, the products for which all consumers have a higher willingness to pay.

The results in Columns [2] and [5] of Panel A in Table 2 indicate that, in over half of the modules represented in the data, the willingness to pay for quality increases with income.\footnote{The full distributions of the \( \gamma_m \) and \( \alpha^1_m \) estimates are depicted in Figures A.2 and A.3 in Appendix E.} Columns [2] and [5] in Panel B show that over half of the estimates for \( \gamma_m \), or 254 and 301 estimates out of 504, are statistically significant in the models that allow for non-homothetic demand for quality and price sensitivity and for non-homothetic demand for quality but not price sensitivity, respectively. Over 75 percent of these statistically significant \( \gamma_m \) estimates are positive in both models. The demand for quality is therefore increasing with income in most grocery sectors.

Columns [3] and [7] of Table 2 summarize the distribution of the estimated values for \( \alpha^1_m \) in each module. In equation (28), one can see that the weight that consumers place on their idiosyncratic utility is increasing in their outside good expenditure, \( Z \), for negative values of \( \alpha^1_m \). Suppose that two consumers draw very high values for \( \varepsilon_{img} \), such that both consumers select to consume product \( g^* \) at the current market price. If the price of product \( g^* \) increases, then the consumer who places a higher weight on his/her idiosyncratic utility draw will be less...
likely to switch to another product for which he/she drew a lower value for $\varepsilon_{img}$ relative to the product’s quality, $\beta_{mg}$. For $\alpha_{1m} < 0$, high-income consumers will place higher weights on their idiosyncratic utility draws and their expenditure shares will, therefore, be less sensitive to price changes. Column [3] of Table 2 shows that, for the majority of modules, high-income consumers are less price sensitive, or $\hat{\alpha}_{1m} < 0$, when you control for the fact that they also have a greater willingness to pay for quality. Column [3] of Panel B in Table 2 indicates that 171, or fewer than half, of the 504 $\alpha_{1m}$ estimates are statistically significant, but over 75 percent of these statistically significant $\alpha_{1m}$ estimates are negative. So, while the evidence that price sensitivity varies with income is less prevalent across modules, the price sensitivity-income pattern is in the expected direction in modules where there is statistically significant variation in the price sensitivity by income.

If we focus instead on the model that allows for non-homothetic price sensitivity but not non-homothetic demand for quality, Column [7] of Table 2 shows that the majority of the $\alpha_{1m}$ estimates, and even the majority of those that are statistically significant, are positive when $\gamma_m$ is constrained to be zero. These estimates may be biased upwards by a correlation between unobserved income-specific product tastes and prices. Consider the model where $\gamma_m$ is restricted to equal zero: $\ln s_{kmgt} - \ln s_{km\bar{g}m} = (\alpha_{0m} + \alpha_{1m}y_k) [(\beta_{mg} - \beta_{m\bar{g}m}) - (\ln p_{kmgt} - \ln p_{km\bar{g}m})] + \nu_{k\bar{g}m,t}$. Here, the error terms include any income-specific product tastes, $\beta_{kmg} - \beta_{km\bar{g}m}$. If the stores at which high-income consumers shop set prices in accordance with these tastes such that $\text{Corr}(\beta_{kmg} - \beta_{km\bar{g}m}, \ln p_{kmgt} - \ln p_{km\bar{g}m}) \neq 0$, then the assumption that $E[Z_t \cdot \nu] = 0$ will be violated. The fact that the $\alpha_{1m}$ 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 includes a term that varies by product and income, $(\beta_{mg} - \beta_{m\bar{g}m}) y_k$, and therefore does not include the full value of $\beta_{kmg} - \beta_{km\bar{g}m}$ in the errors. I do not, therefore, take the positive $\alpha_{1m}$ 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 $\alpha_{1m}$ estimates highlight the difficulty in identifying the non-homotheticity related to price sensitivity in isolation from the non-homotheticity related to product quality.

Despite the results from this one model, the parameter estimates generally show convincing evidence of non-homothetic demand. Specifically, 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.

The upper-level between-module estimation equation (27) yields the elasticity estimates reported in Table 3. $\bar{\sigma} = 1 + \alpha^0 + \alpha^1 \bar{y}$ reflects the elasticity of substitution for the household with the mean income in the markets used in analysis. As expected, products in different modules are
less substitutable than products in the same module, with between module substitution elasticities of between 1.03 and 1.09. The model that allows for non-homotheticity in price sensitivity alone indicates that high-income consumers are more price sensitive than low-income households but, as with the within-module estimates reported above, this pattern reverses to indicate that high-income consumers are less sensitive to cross-module price changes once you control for the fact that they are have a stronger demand for higher quality products and, therefore, the modules containing these brands.

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Parameter ( \bar{\sigma} )</th>
<th>Parameter ( \alpha^1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Homothetic</td>
<td>1.036</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[0.00201]</td>
<td></td>
</tr>
<tr>
<td>2. Non-Homothetic in Price</td>
<td>1.033</td>
<td>0.00043</td>
</tr>
<tr>
<td></td>
<td>[0.00003]</td>
<td>[0.00202]</td>
</tr>
<tr>
<td>3. Non-Homothetic in Quality</td>
<td>1.087</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[0.00945]</td>
<td></td>
</tr>
<tr>
<td>4. Non-Homothetic in Quality and Price</td>
<td>1.047</td>
<td>-0.0018</td>
</tr>
<tr>
<td></td>
<td>[0.00044]</td>
<td>[0.00127]</td>
</tr>
</tbody>
</table>

Note: Standard errors (in brackets) have been adjusted for first-stage measurement error as described in Appendix D.

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 an additional 505 parameters to the model (one \( \alpha^1_m \) for each module, as well as the aggregate cross-module \( \alpha^1 \)). 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.43

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. In Section 6.1 above, I presented estimates of the parameters that govern the within-module product choice for each module \( m \),

---

43This 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).
θ_{1m}, in a separate GMM estimation procedure under the four sets of parameter restrictions corresponding to the four models. For the most flexible version of the model, all elements of \( \theta_{1m} \) are estimated. These include \( \alpha_{0m} \), \( \alpha_{1m} \), \( \gamma_m \), and a relative quality parameter \( (\beta_{mg} - \beta_{m\bar{g}}) \) for each brand represented in the module except for the brand of the base product \( \bar{g} \). For the brand of the base product \( \bar{g} \), \( (\beta_{mg} - \beta_{m\bar{g}}) \) equals zero. Each of the models with parameter restrictions are nested in the full model, which allows for non-homotheticity in both the demand for quality and price sensitivity. Let \( \hat{\theta}_{1m}^M \) denote the estimates for unrestricted component of \( \theta_{1m} \) in model \( M \) and \( \bar{\theta}_{1m}^M \) denote the remaining set of parameters that are restricted to equal zero in model \( M \). All four models have been estimated using the optimal weighting matrix for the full model, which I denote by \( W^*_m \). 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.

The selection criterion minimizes the following GMM-BIC function:

\[
\text{GMM-BIC}^M_m(\hat{\theta}_{1m}^M) = n_m G_m(\hat{\theta}_{1m}^M, \bar{\theta}_{1m}^M)W^*_m G_m(\hat{\theta}_{1m}^M, \bar{\theta}_{1m}^M) - \ln(n_m)(L^*_m - K^M_m)
\]  

(29)

Here, \( G_m(\hat{\theta}_{1m}^M, \bar{\theta}_{1m}^M) \) are the moments for model \( M \) evaluated at the estimated values for free parameters \( \hat{\theta}_{1m}^M \) and zero for the restricted parameters, \( \bar{\theta}_{1m}^M \); \( K^M_m \) 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 \( m \). I evaluate models by calculating the unweighted and sales-weighted share of modules for which that model minimizes the GMM-BIC criterion. The results of this model selection test are presented in Table 4.

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Parameter Restrictions</th>
<th>Unweighted Share</th>
<th>Sales Weighted Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Homothetic</td>
<td>( \gamma_m = 0, \alpha^1_m = 0 )</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>2. Non-Homothetic in Price</td>
<td>( \gamma_m = 0 )</td>
<td>0.15</td>
<td>0.10</td>
</tr>
<tr>
<td>3. Non-Homothetic in Quality</td>
<td>( \alpha^1_m = 0 )</td>
<td>0.47</td>
<td>0.39</td>
</tr>
<tr>
<td>4. Non-Homothetic in Quality and Price</td>
<td>None</td>
<td>0.26</td>
<td>0.36</td>
</tr>
</tbody>
</table>

The model that permits non-homothetic demand for quality, but not for price, is the optimal model for almost half of the modules. These modules represent 39 percent of sample sales. In bilateral model comparisons, the model that accounts for non-homothetic demand for quality had a lower GMM-BIC criterion in modules representing approximately two-thirds of sales when compared to all three of the other models.

\[44\]

The most flexible model has the next-highest share of “winning” modules, representing 36 percent of sample sales. This indicates that the most flexible model performs better in the larger modules. In fact, the most flexible model performs the best in 61 percent of the largest 50 modules by sales, while the least restrictive, or homothetic model, performs the best in 61 percent of the smallest 200 modules by sales. This is most likely related to the number of observations used in the estimation of the largest, relative to the smallest, modules.
The analysis above suggests that the salient form of non-homotheticity in grocery consumption is in the demand for quality. In the analysis below, I therefore study how price indexes that account for this form of non-homotheticity alone vary across cities differently for consumers at different income levels. Any differences between the cross-city price indexes for high- and low-income consumers will reflect cross-city differences in the availability and prices of high-relative to low-quality products. 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 equally.

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, as outlined in Section 5 above. I can now turn to answering the central question in this paper: do grocery costs vary differently across cities for consumers at different income levels?

I address this question using the following baseline regression model:

\[
\ln \hat{P}(P_c, y_k) = \delta_k + \beta_1 y_c + \beta_2 y_k y_c + \epsilon_{kc},
\]

where \(\hat{P}(P_c, y_k)\) denotes the price index for a representative consumer with log income \(y_k\) in city \(c\), \(\delta_k\) is an income-level fixed effect and \(y_c\) is log per capita income in city \(c\).

In the above specification, the elasticity of grocery costs with respect to city income is equal to coefficient on log city income added to the product of log consumer income and by the coefficient on the income interaction term: \(\varepsilon_{P_{c,y}} = \beta_1 + \beta_2 y_k\). The \(\beta_2\) coefficient measures how this elasticity varies with household income. The grocery cost price index, \(\hat{P}(P_c, y_k)\), is calculated using a model that allows for non-homotheticity in the demand for quality, so the \(\varepsilon_{P_{c,y}}\) elasticity will vary with income if the goods and prices available in each city are correlated with the tastes of the incomes of the consumers living there. Suppose, for example, that wealthy cities offer more varieties of high-quality goods at lower prices than poorer cities. If this is the case, 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 who move from poor to 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 $\beta_2$ estimate.\footnote{This regression characterizes an equilibrium relationship and should not be interpreted causally. The results presented here do not indicate whether, for example, price indexes 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.}

The data in Table 1 indicate that market income is correlated with market size; that is, wealthier U.S. cities are larger than poorer U.S. cities with a correlation coefficient of 0.47. Therefore, it is possible that a negative $\beta_2$ 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. If this were the case, a negative $\beta_2$ estimate would support a story in which high-income consumers receive more consumption benefits from living in larger cities than low-income consumers, as opposed to the “preference externalities” story in which high-income consumers receive more consumption benefits from living in wealthier cities and low-income consumers receive more consumption benefits from living in poorer cities. I test between these two theories by adding log population and log population interacted with log household income to the baseline regression.\footnote{Per capita income is not perfectly correlated with population across the sample MSAs: three of the ten largest cities in the sample (Los Angeles, Detroit, and Chicago) and three of the ten smallest (Des Moines, Omaha, and Richmond) have similar per capita incomes, between $35,000 and $40,000. There is, therefore, sufficient variation to identify whether the observed variation in household income- and city-specific grocery costs is related to a story in which high-income households benefit more from city size than low-income consumers, or one in which all consumers benefit more from living in locations with per capita incomes closer to their own.}

The data in Table 1 indicate that market income is correlated with market size; that is, wealthier U.S. cities are larger than poorer U.S. cities with a correlation coefficient of 0.47. Therefore, it is possible that a negative $\beta_2$ 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. If this were the case, a negative $\beta_2$ estimate would support a story in which high-income consumers receive more consumption benefits from living in larger cities than low-income consumers, as opposed to the “preference externalities” story in which high-income consumers receive more consumption benefits from living in wealthier cities and low-income consumers receive more consumption benefits from living in poorer cities. I test between these two theories by adding log population and log population interacted with log household income to the baseline regression.\footnote{This regression characterizes an equilibrium relationship and should not be interpreted causally. The results presented here do not indicate whether, for example, price indexes 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.}

The first column of Table 5 presents the results of the baseline regression. The estimates for $\beta_2$ are negative and statistically significant at the 95 percent level confirming that the elasticity of the price level with respect to per capita income varies across income levels. The magnitude of the $\beta_2$ estimate indicates that this variation is economically significant. A consumer who earns $15,000 a year sees his/her price index rise by around 30 percent for each log unit increase in city per capita income, approximately equivalent to the log difference between San Francisco (per capita income of $54,191) and New Orleans (per capita income of $21,446). On the other hand, the price index of a consumer with a yearly income of $100,000 decreases by around 9 percent for each log unit increase in city per capita income. Therefore, a high-income household experiences a 40 percent greater increase in grocery consumption utility than a low-income household when both move from a poor city to a city with double the per capita income.
Table 5: City-Income Specific Price Index Regressions

<table>
<thead>
<tr>
<th>Dependent Variable: Ln(Price Index for Representative Consumer $k$ in City $c$)</th>
<th>[1]</th>
<th>[2]</th>
<th>[3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Per Capita Income$_c$)</td>
<td>$\beta_1$ 2.412***</td>
<td>-</td>
<td>2.290*</td>
</tr>
<tr>
<td></td>
<td>[0.996]</td>
<td>-</td>
<td>[1.22]</td>
</tr>
<tr>
<td>Ln(Per Capita Income$_c$) * Ln(Household Income$_k$)</td>
<td>$\beta_2$ -0.217**</td>
<td>-</td>
<td>-0.201*</td>
</tr>
<tr>
<td></td>
<td>[0.0915]</td>
<td>-</td>
<td>[0.112]</td>
</tr>
<tr>
<td>Ln(Population$_c$)</td>
<td>$\beta_3$ -</td>
<td>0.28</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>[0.190]</td>
<td>[0.232]</td>
</tr>
<tr>
<td>Ln(Population$_c$)</td>
<td>$\beta_4$ -</td>
<td>-0.027</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>[0.018]</td>
<td>[0.021]</td>
</tr>
</tbody>
</table>

Household Income Fixed Effects: Yes Yes Yes
Observations: 230 230 230
R-Squared: 0.03 0.012 0.036

*** $p<0.01$, ** $p<0.05$, * $p<0.1$; Standard errors in brackets.

The second and third columns of Table 5 show the robustness of this result to city-size controls. When the log price indexes are regressed against these population variables and household income dummies alone, the coefficients on log population and log population interacted with log household income are not statistically significant. When I include these extra variables in the baseline model that controls for city-size effects, the coefficients on the controls remain insignificant. More importantly, the coefficients on log per capita income and log per capita income interacted with log household income are similar in magnitude to the estimates in the baseline model, although their statistical significance has been reduced from the 5 to the 10 percent level.\footnote{Table A.3 in Appendix E shows that these main results only change marginally when based on price indexes that account for non-homotheticities in both consumer’s demand for quality and their price sensitivity.}

Semi-Parametric Estimates

The regression above imposes that the elasticity of the income-specific price index with respect to city income varies linearly with income. There is no reason that this needs to be the case. To obtain non-parametric estimates of these elasticities at different income levels, I estimate the above regression specification but with a household income dummy interacted with per capita city income instead of the household income level interacted with per capita city income:

$$\ln \hat{P}(C, y_k) = \delta + \beta_1 y_c + \beta_2 y_k y_c + \epsilon_{kc},$$

(31)
Figure 6 plots the estimates of the $\beta_{2k}$ elasticity parameters against log household income, $y_k$. These results indicate that there is a non-linear relationship between this elasticity and household income, with the downward slope flattening out at the lower and upper tails of the income distribution. The price index of a consumer who earns $15,000 per year increases by almost 20 percent with each log unit increase in city income, whereas the price index for a consumer who earns $100,000 per year decreases by around 20 percent with each log unit increase in city income.

Figure 6: Elasticity of Grocery Price Index with respect to City Income for Households at Different Size-Adjusted Income Levels

Differentiating between Price and Variety Effects

Taken together, 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. 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 850-household sample for each city, as before, but setting the prices of these products equal to its national average price.
Table 7: City-Income Specific Price Index Regressions

<table>
<thead>
<tr>
<th>Dependent Variable: Ln(Price Index for Representative Consumer $k$ in City $c$)</th>
<th>City-Specific Prices</th>
<th>National Average Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Per Capita Income$_c$)</td>
<td>$\beta_1$</td>
<td>2.412***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.996]</td>
</tr>
<tr>
<td>Ln(Per Capita Income$_c$)</td>
<td>$\beta_2$</td>
<td>-0.217**</td>
</tr>
<tr>
<td>*Ln(Household Income$_k$)</td>
<td></td>
<td>[0.0915]</td>
</tr>
<tr>
<td>Household Income Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>230</td>
<td>230</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.03</td>
<td>0.098</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1; Standard errors in brackets.

The second column of Table 7 reports the estimates of the baseline regression model run using these fixed-price indexes as the dependent variable. The first column replicates the results baseline results from Table 5. We observe that the coefficient on the interaction between per capita income and household income is more negative and more statistically significant at a higher level when the fixed-price indexes are used as the dependent variable. High-income households would find wealthy cities to be even less expensive than poor cities, relative to low-income households, if products were sold in both locations at their national average price. This suggests that the entire difference between how high- and low-income households perceive the relative costs to vary across cities is due to variety differences. The products that high-income consumers prefer to consume are sold at higher prices in wealthy cities than they are in poor cities, but high-income consumers are more than compensated for this price difference by the fact that more of these products are available to them in these locations.

6.4 Comparison with Homothetic Index

I now turn to addressing the extent to which homotheticity biases the estimates of cross-city price indexes for consumers at different income levels. If we assume that preferences are homothetic such that all households get the same utility from the consumption baskets available in one market relative to another, we only need one homothetic price index to compare the utility that households get in one city relative to another. By allowing preferences to be non-homothetic, I allow households at different income levels to get different relative utilities from the consumption baskets available in different locations and, therefore, calculate a different price index to measure these relative utilities for each income-level.

The analysis above has shown that there is economically significant variation in how these
non-homothetic price indexes vary across cities for consumers at different income levels. A homothetic price index captures none of this variation, but it may match the cross-city variation in prices for consumers at some income levels better than others. To consider this question, I first calculate a homothetic price index for each city using the parameter estimates for the model that does not permit either the demand for quality or the price sensitivity of a household to vary with income. Panel A of Table 9 I compare these cross-city homothetic price indexes to the income-specific cross-city price indexes calculated using the parameter estimates for the selected model, which permits the demand for quality to vary with income. The homothetic price index is highly correlated with the non-homothetic price indexes calculated for households earning below $70,000 per year. The correlation between the homothetic price index and the non-homothetic indexes is highest with a coefficient of 0.97 for households earning around $50,000 per year. This indicates that the homothetic price index does a good job at predicting the cities in which low- and middle-income households will gain the most, and the least, from the grocery consumption bundles available there. The correlation coefficient drops to 0.21 for households earning around $80,000 per year and is negative for households earning around $150,000 per year. The homothetic price index, therefore, does a poor job of predicting which cities high-income households find the most and the least expensive.

Panel B of Table 9 further illustrates these facts. While the homothetic model does not perfectly predict the rankings of cities for households at any income level, it performs very poorly in predicting the most and least expensive cities for high-income households. The homothetic model predicts that Chicago, San Antonio, Sacramento, and San Francisco, are among the six most expensive cities for purchasing groceries. The non-homothetic model, however, predicts that these four cities are among the six cheapest for a household earning $150,000 per year. Conversely, the homothetic model predicts that Atlanta, Detroit, and Columbus are among the five cheapest cities for purchasing groceries, while the non-homothetic model predicts that these cities are among the five most expensive cities for households earning either $93,000 or $150,000 per year.48

6.5 Robustness Checks

Robustness to Demographic Sampling Bias

The prices and products included in the indexes for each city are based on the purchases of a random sample of 850 households Nielsen sample in each city with 850 or more households during 2005. As discussed earlier, Nielsen samples a demographically representative sets of households in each city. This implies that, even holding the number of households constant

48Tables A.5 and A.6 in Appendix E provide the ranks and levels of the homothetic and non-homothetic price indexes for all 23 cities in the price index sample.
Table 9: Comparison of Grocery Costs Calculated Using Homothetic and Non-Homothetic Models

Panel A: Correlation Between Homothetic and Non-Homothetic Price Indexes

<table>
<thead>
<tr>
<th>Household Income</th>
<th>$16,896</th>
<th>$26,715</th>
<th>$35,715</th>
<th>$41,526</th>
<th>$53,103</th>
<th>$60,442</th>
<th>$64,805</th>
<th>$82,576</th>
<th>$93,411</th>
<th>$146,566</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.83</td>
<td>0.88</td>
<td>0.94</td>
<td>0.96</td>
<td>0.97</td>
<td>0.92</td>
<td>0.83</td>
<td>0.21</td>
<td>0.03</td>
<td>-0.11</td>
</tr>
</tbody>
</table>

Panel B: City Ranks (Least Expensive to Most Expensive)

Six Least Expensive Cities According to Homothetic Model

<table>
<thead>
<tr>
<th>Market</th>
<th>Homothetic</th>
<th>Non-Homothetic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$16,896</td>
<td>$26,715</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>DC-Baltimore</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Atlanta</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Detroit</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Columbus</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Tampa</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>

Six Most Expensive Cities According to Homothetic Model

<table>
<thead>
<tr>
<th>Market</th>
<th>Homothetic</th>
<th>Non-Homothetic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$16,896</td>
<td>$26,715</td>
</tr>
<tr>
<td>Chicago</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>San Antonio</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>Sacramento</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>San Francisco</td>
<td>23</td>
<td>23</td>
</tr>
</tbody>
</table>
across cities, we will observe more purchases of products that high-income households prefer to consume in high-income cities. To check whether the results above are mechanically driven by this sample bias, I re-calculate the price indexes using a stratified sample of households for each city, including 190 randomly-sampled households from each tercile of the income distribution. This limits the number of cities included in the analysis to 22, since San Francisco has fewer than 190 low-income households.

Figure 7 indicates that the magnitude and direction of the results presented above is robust to stratified sampling. Low-income households find the cost of the basket of prices and products observed in the stratified sample for high-income cities to be approximately 15 percent higher than it is for low-income cities, while high-income households find these costs to be approximately 5 percent lower. While the magnitude of these price index differences is lower with the stratified sample than with the demographically-representative sample, the gap between how high- and low-income households perceive the relative costs between high- and low-income cities is still large at 20 percent and the main qualitative result still holds: low-income households find wealthy cities to be more expensive, while high-income households find them to be less expensive.

Figure 7: Elasticity of Grocery Price Index with respect to City Income for Households at Different Size-Adjusted Income Levels Using a Stratified Sample of Households

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 below will 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 assumption made 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 household size. Empirically, any correlations that do exist between consumer demand and demographics will be accounted for empirically only to the extent that these characteristics are correlated with size-adjusted household income, in which case these patterns will be attributed to non-homotheticities, biasing the estimates of the model parameters that govern them. These omitted variable biases will only carry through to bias the cross-city price indexes if demographics vary across wealthy and poor cities differently for households earning different incomes. I do not expect these biases to be large as there is no evidence for these correlations in the data.49

49Suppose, for example, that high-income households are more educated and that demand for quality is correlated with education, and not income. The model and estimates presented below will attribute this correlation to non-homotheticities. The price index estimates will still be a valid representation of the relative price levels faced by households at different income levels across cities insofar as education levels do not vary across cities differently for households in different income groups. Table A.7 in Appendix E supports this assumption for a range of characteristics, including education. Sample households with the same size-adjusted income do not in general have systematically different demographic characteristics in wealthy, relative to poor, cities and, even in the one case that a pattern does exist (households are more likely to be married in wealthier cities), it is no stronger for
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 their utilities across these markets. The results in this paper indicate that this is indeed the case: high-income households face greater consumption gains from moving to high per capita income markets than do low-income households.

I show that high-income households face 20 percent lower grocery costs in wealthy cities than in poor cities, while low-income households face 20 percent 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 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. The gap in relative grocery costs is large enough to suggest that this analysis is worthwhile. Even if we assume that preferences are homothetic within each of the households other consumption areas, the difference in relative grocery costs alone implies an economically-significant 2.4 percent gap between the aggregate living costs faced by high-income households in wealthy, relative to poor, cities and those faced by low-income households. High-income households, who spend around 2 percent of their annual income on groceries, would face 0.4 percent lower living costs in wealthy cities, whereas low-income households, who spend around 10 percent of their annual income on groceries, would face 2 percent higher living costs in these locations.

The main goal of this paper is to measure how living costs vary across cities differently for consumers earning different incomes. In doing so, however, I also provide a methodological framework that could be applied more generally, for example, in analyzing how consumption costs vary differently across countries and over time for consumers earning different incomes. Though the detailed household-level data used in this paper might not be available in these contexts, it is conceivable that simulation techniques could be used to identify the parameters of the model presented here using aggregate market- or country-level data. Where I have provided some sense of the large differences in cross-city price indexes across income levels, these extensions could shed light on how other key economic statistics, such as purchasing power parity, inflation, and the consumption gains from trade, vary with income.
References


Appendices for Online Publication

A Non-Homotheticity Condition

Suppose that consumers select grocery consumption quantities, \( Q = \{q_{mg} \}_{g \in G_m} \), and non-grocery expenditure, \( Z \), by maximizing:

\[
 f(U_iG(Q, Z), Z) \quad \text{subject to} \quad \sum_{m \in M} \sum_{g \in G_m} p_{mg} q_{mg} + Z \leq Y_i, \; q_{mg} \geq 0 \; \forall \; mg \in G \quad (A.1)
\]

I break this problem into two parts, first solving for the consumer’s optimal grocery consumption quantities conditional on their non-grocery expenditure \( Z \):

\[
 \max_{Q, Z} U_iG(Q, Z) = \left\{ \sum_{m \in M} \left( \sum_{g \in G_m} q_{mg} \exp(\gamma_m(Z)\beta_m + \mu_m(Z)\epsilon_{mg}) \right)^{\frac{\sigma(Z)-1}{\sigma(Z)}} \right\}^{\frac{1}{\sigma(Z)-1}} \quad \text{subject to} \quad \sum_{m \in M} \sum_{g \in G_m} p_{mg} q_{mg} \leq Y_i - Z, \; q_{mg} \geq 0 \; \forall \; mg \in G \quad (A.2)
\]

where \( \gamma_m(Z) = (1 + \gamma_m \ln Z) \), \( \mu_m(Z) = \frac{1}{\alpha_m + \alpha_1 \ln Z} \), and \( \sigma(Z) = 1 + \alpha^0 + \alpha^1 \ln Z \). Equations (9), (12), and (13) define the optimal grocery bundle, \( Q^*(Z) = \{q_{mg}^*\}_{m \in M} \) and can be summarized as follows:

\[
 q_{mg}^*(Z) = \begin{cases} 
 (Y_i - Z)\frac{\tilde{p}_{mg}}{P_i(Z)}^{\frac{\sigma(Z)-1}{1-\sigma(Z)}} / p_{mg} & \text{if } g = \arg \max_{g \in G_m} \tilde{p}_{mg} \\
 0 & \text{otherwise}
\end{cases}
\]

where

\[
 P_i(Z) = \left[ \sum_{m \in M} \max_{g \in G_m} \tilde{p}_{mg} \right]^{\frac{\sigma(Z)-1}{1-\sigma(Z)}}
\]

and

\[
 \tilde{p}_{mg} = \frac{\exp(\gamma_m(Z)\beta_m + \mu_m(Z)\epsilon_{mg})}{p_{mg}}
\]
Plugging this solution into \( U_{iG}(Q, Z) \) yields the consumer’s indirect utility from grocery consumption, conditional on their non-grocery expenditure:

\[
\tilde{U}_{iG}(Z) = U_{iG}(Q^*(Z), Z)
\]

\[
= \left\{ \sum_{m \in M} \left[ (Y_i - Z) \left( \frac{\tilde{p}_{img}}{P_i(Z)^{1-\sigma(Z)}} \right) \right] \left[ g = \arg\max_{g \in G_m} \tilde{p}_{img} \right] \right\}^{\frac{\sigma(Z)-1}{\sigma(Z)}}
\]

\[
= \frac{Y_i - Z}{P_i(Z)^{1-\sigma(Z)}} \left\{ \sum_{m \in M} \left[ \frac{\tilde{p}_{img}}{P_i(Z)^{1-\sigma(Z)}} \right] \left[ g = \arg\max_{g \in G_m} \tilde{p}_{img} \right] \right\}^{\frac{\sigma(Z)-1}{\sigma(Z)}}
\]

\[
= \frac{Y_i - Z}{P_i(Z)^{1-\sigma(Z)}} \left\{ \sum_{m \in M} \left( \frac{\max_{g \in G_m} \tilde{p}_{img}}{P_i(Z)^{1-\sigma(Z)}} \right)^{\sigma(Z)-1} \right\}^{\frac{\sigma(Z)-1}{\sigma(Z)}}
\]

\[
= \frac{Y_i - Z}{P_i(Z)^{1-\sigma(Z)}} P_i(Z)^{\sigma(Z)}
\]

\[
= \frac{Y_i - Z}{P_i(Z)}
\]  

(A.3)

We can now express problem (A.1) to be a choice over one variable, \( Z \):

\[
\max_Z f(\tilde{U}_{iG}(Z), Z)
\]  

(A.4)

The first order condition to the utility maximization problem defined in problem (A.4) with respect to \( Z \) is:

\[
f_1(\tilde{U}_{iG}(Z), Z) \frac{\partial \tilde{U}_{iG}(Z)}{\partial Z} + f_2(\tilde{U}_{iG}(Z), Z) = 0
\]

Substituting the maximized grocery expenditure conditional on \( Z, \tilde{U}_{iG}(Z) \), from equation (A.3) into this first order condition yields a function that implicitly defines the optimal non-grocery expenditure, \( Z_i \), in terms of household income, \( Y_i \), the consumer’s idiosyncratic utility draws, \( \varepsilon_i \), and model parameters:

\[
Y_i = \frac{Z - P_i(Z)}{P_i(Z)^{1-\sigma(Z)}} + \frac{f_2(\tilde{U}_{iG}(Z), Z) P_i(Z)^2}{f_1(\tilde{U}_{iG}(Z), Z) P_i(Z)}
\]

Taking the derivative of income with respect to outside good expenditure, \( Z \), we can see that the outside good will be normal if the price vector and aggregate utility function are such that:

\[
\frac{\partial}{\partial Z} \left[ \frac{P_i(Z)}{P_i(Z)^{1-\sigma(Z)}} + \frac{f_2(\tilde{U}_{iG}(Z), Z) P_i(Z)^2}{f_1(\tilde{U}_{iG}(Z), Z) P_i(Z)} \right] < 1
\]

It is computationally infeasible to show that this condition holds generally (there will be a different price index \( P_i(Z) \) for each of universe of potential price vectors), but I can show that it holds in the data by simply demonstrating that non-grocery expenditures are increasing in household income. I annualize the observed grocery expenditure for each household and measure annual non-grocery expenditures as the difference between the mid-point of each household’s reported income category and the household’s annual grocery expenditures. Figure A.1 plots the resulting non-grocery expenditure variable for households in each reported income category,
demonstrating a clear increasing relationship between non-grocery expenditures and income. After controlling for household demographics, the elasticity of observed non-grocery expenditures, $Z_i$, with respect to household income, $Y_i$, is 1.05 with a standard error of 0.0003.

Figure A.1: Household Grocery Expenditure by Income Category

Notes: 1. Household grocery expenditures are calculated by subtracting each household’s annualized observed grocery expenditures from their annual income, inferred from their reported annual income category as $150,000 for the above $100,000 category and the mid-point of the reported income category bounds for all other income categories. 2. Boxes represent the inter-quartile range of household grocery expenditures. The dots outside the whiskers represent outlier observations.

B Derivations

B.1 Within-Module Consumption Decision

Consumer $i$, spending $Z$ on the outside good, chooses how to allocate expenditures between products within a module $m$ conditional on their expenditure in that module, $w_m$, to maximize

$$u_{im}(w_m, Z) = \sum_{g \in G_m} q_{mg} \exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img})$$

subject to the module-level budget constraint, $\sum_{m \in M} \sum_{g \in G_m} p_{mg}q_{mg} \leq w_m$, and non-negativity constraints $q_{mg} \geq 0$.

Recall that the additive log-logit functional form implies that consumers optimally purchase a positive quantity only one product in a module. This product maximizes their marginal utility of expenditure in a module conditional
on their outside good expenditure: \(^{50}\)

\[
g_{im}^*(Z) = \arg \max_{g \in G_m} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img})}{p_{mg}}
\]  

(A.5)

Since all of a consumer’s module expenditure, \(w_m\), is allocated to this optimal product, \(g_{im}^*\), the consumer’s optimal module bundle, \(Q_{im}^*(w_m, Z)\), can be written as:

\[
Q_{im}^*(w_m, Z) = \left( q_{im1}^*(w_m, Z), \ldots, q_{imG_m}^*(w_m, Z) \right)
\]

where \(q_{img}^*(w_m) = \begin{cases} w_m/p_{mg} & \text{if } g = \arg \max_{g \in G_m} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img})}{p_{mg}} \\ 0 & \text{otherwise} \end{cases} \)  

(A.6)

That is, a consumer \(i\) optimally consumes as much of their optimal product, \(g_{im}^*(Z)\), as their module expenditure, \(w_m\), will afford them and zero of any other product in the module.

### B.2 Across-Module Consumption Decision

Consumer \(i\), spending \(Z\) on the outside good, chooses how to allocate expenditures between modules by selecting \(w_1, \ldots, w_M\) to maximize

\[
U_i(w_1, \ldots, w_M) = \left\{ \sum_{m \in M} \left[ \tilde{u}_{im}(w_m, Z) \right]^{\rho_i} \right\}^{\frac{1}{\rho_i}} = \left\{ \sum_{m \in M} \left[ w_m \max_{g \in G_m} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img})}{p_{mg}} \right]^{\frac{\sigma(z)-1}{\sigma(z)}} \right\}^{\frac{\sigma(z)}{\sigma(z)-1}}
\]

subject to

\[
\sum_{m \in M} w_m \leq Y_i - Z
\]

We simplify the expression for the target utility function by denoting consumer \(i\)’s marginal utility from expenditure in module \(m\) as the inverse of \(A_{im}\):

\[
\max_{g \in G_m} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img})}{p_{mg}} = \frac{1}{A_{im}}
\]

(A.7)

The within-module allocation decision now simplifies to:

\[
w_i^*(Z) = \left( w_{i1}^*(Z), \ldots, w_{iM}^*(Z) \right) = \arg \max \left\{ \sum_{m \in M} \frac{w_m}{A_{im}} \right\} \left\{ \sum_{m \in M} w_m \leq Y_i - Z \right\}
\]

(A.8)

The utility function over module expenditures is concave in module expenditure for each module \(m\). Therefore, there will be an interior solution to the maximization problem and it can be solved using the first order conditions

\(^{50}\text{Note that the marginal utility of expenditure in a module and, therefore, the optimal product choice, } g_{im}^*, \text{ depends on a consumer’s outside good expenditure, } Z, \text{ but is independent of their module expenditure, } w_m.\)
with respect to expenditure in each module \( m \). The first order condition for each module \( m \) is:

\[
\frac{\partial U_i(w_1, \ldots, w_M)}{\partial w_m} = \left\{ \sum_{m \in M} \left[ \frac{w_m}{A_{im}} \right]^{\frac{\sigma(Z)-1}{\sigma(Z)}} \right\}^{\frac{1}{1-\sigma(Z)}} \cdot \frac{1}{A_{im}} \left[ \frac{w_m}{A_{im}} \right]^{\frac{1}{\sigma(Z)}} = \lambda
\]

where \( \lambda \) is the marginal utility of expenditure. This implies that the marginal utility of expenditure must be equal across modules. We use this equality across two modules, \( m \) and \( m' \), to solve for the optimal expenditure in module \( m' \):

\[
\left\{ \sum_{m \in M} \left[ \frac{w_m}{A_{im}} \right]^{\frac{\sigma(Z)-1}{\sigma(Z)}} \right\}^{\frac{1}{1-\sigma(Z)}} \cdot \frac{1}{A_{im'}} \left[ \frac{w_m}{A_{im'}} \right]^{\frac{1}{\sigma(Z)}} = \left\{ \sum_{m \in M} \left[ \frac{w_m}{A_{im}} \right]^{\frac{\sigma(Z)-1}{\sigma(Z)}} \right\}^{\frac{1}{1-\sigma(Z)}} \cdot \frac{1}{A_{im}} \left[ \frac{w_m}{A_{im}} \right]^{\frac{1}{\sigma(Z)}}
\]

\[
w_{m'} = w_m \left[ \frac{A_{im'}}{A_{im}} \right]^{1-\sigma(Z)}
\]

Imposing the budget constraint, \( \sum_{m \in M} w_{m'} = \sum_{m \in M} w_m \leq Y_i - Z \), yields an expression for \( w_m \) in terms of total expenditure, \( Y_i - Z \), and an index of the \( A_{im} \) terms:

\[
Y_i - Z = \sum_{m' \in \mathcal{M}} w_{m'}
\]

\[
Y_i - Z = w_m \frac{A_{im}}{A_{im'}} \sum_{m' \in \mathcal{M}} [A_{im'}]^{1-\sigma(Z)}
\]

\[
w_m = \sum_{m' \in \mathcal{M}} [A_{im'}]^{1-\sigma(Z)} (Y_i - Z)
\]

The solution to problem (A.8) is, therefore,

\[
w^*_i(Z) = (w^*_1(Z), \ldots, w^*_M(Z)) \quad \text{where} \quad w^*_m = \frac{A_{im}^{1-\sigma(Z)}}{P_i^{1-\sigma(Z)}} (Y_i - Z) \quad \forall m \in \mathcal{M}
\]

where \( P_i(Z) \) is a CES price index over \( A_{im} \) for all modules \( m \in \mathcal{M} \) defined as:

\[
P_i(Z) = \left[ \sum_{m \in \mathcal{M}} A_{im}^{1-\sigma(Z)} \right]^{\frac{1}{1-\sigma(Z)}}
\]

Substituting from equation (A.7) for \( A_{img} \) yields consumer \( i \)'s optimal module expenditure vector, \( w^*_i(Z) \), as a function of total grocery expenditures, prices, and model parameters:

\[
w^*_i(Z) = (w^*_1(Z), \ldots, w^*_M(Z)) \quad \text{where} \quad w^*_m = (Y_i - Z) \left[ \max_{g \in \mathcal{G}_m} \frac{\exp(\gamma_{mg} Z_{mg} + \mu_m(Z) Z_{img})}{p_{img}} \right]^{\sigma(Z)-1}
\]

\[
w^*_m = \left[ \frac{\max_{g \in \mathcal{G}_m} \frac{\exp(\gamma_{mg} Z_{mg} + \mu_m(Z) Z_{img})}{p_{img}}}{P_i(Z)^{1-\sigma(Z)}} \right]^{1/\sigma(Z)}
\]

53
\[ P_i(Z) = \left[ \sum_{m \in M} \left( \max_{g \in G_m} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img})}{p_{mg}} \right)^{\sigma(Z)-1} \right]^{\frac{1}{\sigma(Z)}} \]

### B.3 Within-Module Market Expenditure Shares

Equation (9) states that:

\[ Q_{im}^*(w_m, Z) = (q_{im1}^*(w_m, Z), \ldots, q_{imG_m}^*(w_m, Z)) \text{ where } q_{img}^*(w_m, Z) = \begin{cases} \frac{w_m}{p_{mg}} & \text{if } g = \arg \max_{g \in G_m} \tilde{p}_{img} \\ 0 & \text{otherwise} \end{cases} \]

where \( \tilde{p}_{img} = \exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img}) \). If we rewrite consumer \( i \)'s optimal consumption quantity using an indicator function to identify which product is selected by the consumer, consumer \( i \)'s optimal consumption quantity of product \( g \) in module \( m \) is:

\[ q_{img}^*(w_m, Z) = \frac{w_m}{p_{mg}} \mathbb{1} \left[ g = \arg \max_{g \in G_m} \tilde{p}_{img} \right] \]

We can use this definition to derive consumer \( i \)'s expenditure on product \( g \) in module \( m \), conditional on their outside good expenditure \( Z \) and the vector of module prices they face, \( P_m \):

\[ s_{img|m}(Z, P_m) = \mathbb{1} \left[ g = \arg \max_{g \in G_m} \tilde{p}_{img} \right] \]

The expected value of this expenditure share is derived by integrating over the idiosyncratic utilities in module \( m \), \( \epsilon_{im} \):

\[ \mathbb{E}_{\epsilon}[s_{img|m}(Z, P_m)] = \mathbb{E}_{\epsilon} \left[ \mathbb{1} \left[ g = \arg \max_{g \in G_m} \tilde{p}_{img} \right] \right] = P_r \left[ \tilde{p}_{img} \geq \tilde{p}_{img'}, \forall g' \in G_m \right] = P_r \left[ \epsilon_{img} - \epsilon_{img'} \geq \frac{\gamma_m(Z)(\beta_{mg} - \beta_{mg'}) - (\ln p_{mg} - \ln p_{mg'})}{\mu_m(Z)} \right] = \sum_{g' \in G_m} \tilde{p}_{img'} \]

The final equality holds because the idiosyncratic utilities, \( \epsilon_{im} \), are iid draws from a type I extreme value distribution. Imposing the parametric forms for \( \gamma_m(Z) = (1 + \gamma_m \ln Z) \) and \( \mu_m(Z) = (\alpha_m^0 + \alpha_m^1 \ln Z)^{-1} \) from equations (4) and (5), respectively, ensures that the consumer's expected expenditure share is common with other consumers with the same income that face the same product prices:

\[ \mathbb{E}_{\epsilon}[s_{img|m}(Z, P_m)] = \frac{\exp[(\alpha_m^0 + \alpha_m^1 \ln Z)((1 + \gamma_m \ln Z)\beta_{mg} - \ln p_{mg})]}{\sum_{g' \in G_m} \left( \exp[(\alpha_m^0 + \alpha_m^1 \ln Z)((1 + \gamma_m \ln Z)\beta_{mg'} - \ln p_{mg'})] \right)} \]
I interpret the expected expenditure share function derived above as the expected share of expenditure that a group of households with the same outside good expenditure, $Z$, facing identical prices for products in module $m$ spend on product $g$. If the group of households is in the same market, then this expected expenditure share will be the income-specific market share of product $g$ in module $m$, which I denote by $s_{mg|m}(Z, P_m)$. $s_{mg|m}(Z, P_m)$ is the share of expenditure that a group of households with the outside good expenditure, $Z$, and facing a common vector of module prices, $P_m$:

$$ s_{mg|m}(Z, P_m) = E_c[s_{img|m}(Z, P_m)] = \frac{\exp[(\alpha_m^0 + \alpha_m^1 \ln Z)(\beta_{mg}(1 + \gamma_m \ln Z) - \ln p_{mg})]}{\exp[(\alpha_m^0 + \alpha_m^1 \ln Z)(\beta_{mg}(1 + \gamma_m \ln Z) - \ln p_{mg^*})]} $$

Dividing this market share for product $g$ in module $m$ by the market share for a fixed product $\bar{g}_m$ in the same module $m$ results in a relative market share that depends only on model parameters, consumer income, and the prices of product $g$ and $\bar{g}_m$:

$$ \frac{s_{mg|m}(Z, P_m)}{s_{\bar{g}g|m}(Z, P_m)} = \frac{\exp[(\alpha_m^0 + \alpha_m^1 \ln Z)(\beta_{mg}(1 + \gamma_m \ln Z) - \ln p_{mg})]}{\exp[(\alpha_m^0 + \alpha_m^1 \ln Z)(\beta_{mg}(1 + \gamma_m \ln Z) - \ln p_{mg^*})]} $$

I linearize the relative expenditure share equation by taking the log of both sides:

$$ \ln(s_{mg|m}(Z, P_m)) - \ln(s_{\bar{g}g|m}(Z, P_m)) = (\alpha_m^0 + \alpha_m^1 \ln Z) [(\beta_{mg} - \beta_{mg^*})(1 + \gamma_m \ln Z) - (\ln p_{mg} - \ln p_{mg^*})] $$

Equation (A.9) defines the expected within-module expenditure share of a set of households with outside good expenditure $Z$ facing prices $p_{mg}$ and $p_{mg^*}$ on product $g$ in module $m$ relative to product $\bar{g}_m$ in the same module $m$ in terms of parameters $\alpha_m, \gamma_m$, and $(\beta_{mg} - \beta_{mg^*})$. This equation is used to calculate moments for each product $g \neq \bar{g}_m$ in each module $m$, that are in turn used to estimate all of the $\alpha_m$ and $\gamma_m$ parameters, as well as each $\beta_{mg}$ parameter relative to $\beta_{mg^*}$, i.e. $(\beta_{mg} - \beta_{mg^*})_{g \in G_m}$.

### B.4 Between-Module Relative Market Expenditure Shares

I now want to generate a similar estimation equation that can be used to identify $\alpha_m^0$, $\alpha_m^1$, and $(\beta_{mg^*})_{g \in G_m}$ using data on module-level income-specific market shares. Equations (12) and (13) together characterize the optimal cross-module expenditure allocation for consumer $i$ conditional on this consumer’s idiosyncratic utility draws for each product in each module. These equations are:

$$ w_i^*(Z, P) = (w_{i1}(Z, P), ..., w_{iM}(Z, P)) \text{ where } w_{im}^* = (Y_i - Z) \left[ \max_{g \in G_m} \tilde{p}_{img} \right]^{\sigma(Z)-1} \left( P_i(Z) \right)^{1-\sigma(Z)} $$

$$ P_i(Z, P) = \left[ \sum_{m \in M} \left( \max_{g \in G_m} \tilde{p}_{img} \right)^{\sigma(Z)-1} \right]^{-1/\sigma(Z)} $
expression. I take the log of this relative share expression to linearize the equation:

\[ \ln(\text{relative share}) \]

Since \( \ln(\) taking logs and differencing to generate the estimation equation (A.9), as it demonstrates the relationship between the consumer’s expected product expenditure share over the idiosyncratic errors, \( E[s_{img}(Z, \mathbb{P}_m)] \), to derive an expression for the market share of each product. I then divide these market shares by the market share of a module specific base product and taking logs to linearize the equation. I change the order of this procedure when deriving the between-module relative market share equation, i.e. difference and take the log of the individual’s expenditure shares before taking the expectation of these terms over the idiosyncratic errors. The reason for this reordering is that the consumer’s module expenditure shares include a term, \( P_i \), that depends non-linearly on all of the consumer’s idiosyncratic utility draws. This term is common to all of the consumer’s module shares, and thus drops out of the consumer’s relative module expenditure shares, so that these relative shares are functions of the consumer’s idiosyncratic utility draws in the two relevant modules. The log of this relative module expenditure share term is additive in terms that depend on the consumer’s idiosyncratic utility draws in only one module at a time; that is, a term that depends on the consumer’s idiosyncratic utility draws in module \( m \) and a term that depends on the consumer’s idiosyncratic utility draws in the base module \( \bar{m} \). This makes the expectation of the consumer’s log expenditure share in module \( m \) relative to module \( \bar{m} \) easier to derive than the expectation of the consumer’s expenditure share for a single module \( m \).

I now generate the relative module market shares. As discussed above, I first divide consumer \( i \)’s module expenditure share, \( s_{im}(Z, \mathbb{P}) \), by his/her expenditure share in some fixed base module \( \bar{m} \):

\[
s_{im}(Z, \mathbb{P}) = \frac{\bar{w}_{im}(Z)}{\bar{Y}_i - Z} = \left[ \max_{g \in \mathbb{G}_m} \tilde{P}_{img} \right]_{\sigma(Z)-1}^{\sigma(Z)-1} \left[ \max_{g \in \mathbb{G}_m} \tilde{P}_{m\bar{m}g} \right]_{\sigma(Z)-1}^{\sigma(Z)-1}
\]

Since \( P_i \) does not vary across modules for a given consumer \( i \), it drops out of the relative module expenditure share expression. I take the log of this relative share expression to linearize the equation:

\[
\ln s_{im}(Z, \mathbb{P}) - \ln s_{im}(Z, \mathbb{P}) = (\sigma(Z) - 1) \ln \left( \max_{g \in \mathbb{G}_m} \tilde{P}_{img} \right) - (\sigma(Z) - 1) \ln \left( \max_{g \in \mathbb{G}_m} \tilde{P}_{m\bar{m}g} \right),
\]

52 The order of the expectation, differencing, and log operations does not make a difference to the relative market share equation in the within-module case, that is:

\[
\ln(s_{mg|m}(Z, \mathbb{P}_m)) - \ln(s_{mg|m}(Z, \mathbb{P}_m)) = \ln \left[ E[\ln(s_{img|m}(Z, \mathbb{P}_m))] / E[\ln(s_{img|m}(Z, \mathbb{P}_m))] \right] = E[\ln(s_{img|m}(Z, \mathbb{P}_m)) - \ln(s_{img|m}(Z, \mathbb{P}_m))] = (\alpha_m^0 + \alpha_m^1 \ln Z) (1 + \gamma_m \ln Z) - (\ln P_{mg} - \ln P_{m\bar{m}})
\]

I derive the expression for the Z-specific market share of product \( g \), \( s_{mg|m}(Z, \mathbb{P}_m) = E[\ln(s_{img|m}(Z, \mathbb{P}_m))] \), before taking logs and differencing to generate the estimation equation (A.9), as it demonstrates the relationship between the term on the left-hand side of this equation, \( \ln(s_{mg|m}(Z, \mathbb{P}_m)) - \ln(s_{mg|m}(Z, \mathbb{P}_m)) \), and its value in the data: the difference between the log of the expenditure consumers spending \( Z \) on the outside good in a given market on product \( g \) relative to the log of their expenditure on the base product \( \bar{g} \) or, more succinctly, the log difference between the Z-specific market shares on products \( g \) and \( \bar{g} \).
This equation is a linear function of two terms, the first of which depends on the consumer’s idiosyncratic utility draws in only module \( m \) and the second of which depends on the consumer’s idiosyncratic utility draws in only module \( \bar{m} \). The expectation of the log difference between the consumer’s module expenditure shares can be split into the difference between two expected values:

\[
E_c [\ln s_{im}(Z, P) - \ln s_{im}(Z, \mathbb{P})] = (\sigma(Z) - 1) \left\{ E_c \left[ \ln \left( \max_{g \in G_m} \tilde{p}_{img} \right) \right] - E_c \left[ \ln \left( \max_{g \in G_{\bar{m}}} \tilde{p}_{img} \right) \right] \right\} \tag{A.10}
\]

Consider the two expectation terms in equation (A.10). Both take the same form, and thus I only solve for the first expectation:

\[
E_c \left[ \ln \left( \max_{g \in G_m} \tilde{p}_{img} \right) \right] \tag{A.11}
\]

The expectation term defined in equation (A.11) is the expected value of the log of a maximum. Since the log is a monotonically increasing function, we can switch the order of the log and maximum functions inside the expectation and linearize to yield:

\[
E_c \left[ \ln \left( \max_{g \in G_m} \tilde{p}_{img} \right) \right] = E_c \left[ \ln \left( \max_{g \in G_m} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img})}{p_{mg}} \right) \right] \\
= E_c \left[ \max_{g \in G_m} \ln \left( \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img})}{p_{mg}} \right) \right] \\
= E_c \left[ \max_{g \in G_m} \gamma_m(Z)\beta_{mg} - \ln p_{mg} + \mu_m(Z)\epsilon_{img} \right] \\
= \mu_m(Z)E_c \left[ \max_{g \in G_m} (\gamma_m(Z)\beta_{mg} - \ln p_{mg})/\mu_m(Z) + \epsilon_{img} \right] \tag{A.12}
\]

De Palma and Kilani (2007) show that, for an additive random utility model with \( u_i = v_i + \epsilon_i, i = 1, \ldots, n \) and \( \epsilon_i \sim F(x) \) a continuous CDF with finite expectation, the expected maximum utility is:

\[
E[E_c[\max_i u_i + \epsilon_i] = \int_{-\infty}^{\infty} zd\phi(z) \text{ where } \phi(z) = Pr[\max_k \nu_k \leq z] = \prod_{k=1}^{n} F(z - \nu_k)
\]

Since the expectation in equation (A.12) takes the form \( E_c[\max_i u_{img} + \epsilon_{img}] \), with \( u_{img} = (\gamma_m(Z)\beta_{mg} - \ln p_{mg})/\mu_m(Z) \), and since I have assumed that \( \epsilon_{img} \sim F(x) \) for \( F(x) = \exp(-\exp(-x)) \), I can use the de Palma and Kilani (2007) result to solve for the expectation as follows, dropping the \( i \) and \( m \) subscripts for the notational convenience:

\[
E_c \left[ \max_{g \in G_m} \nu_g + \epsilon_g \right] = \int_{-\infty}^{\infty} zd\phi(z) \\
= \int_{-\infty}^{\infty} zd \left[ \prod_{g=1}^{G_m} \exp(-\exp(v_g - z)) \right] \\
= \int_{-\infty}^{\infty} z \left[ \sum_{g=1}^{G_m} \exp(v_g - z) \right] \exp \left( \sum_{g=1}^{G_m} -\exp(v_g - z) \right) dz
\]
Let $V = \ln \left[ \sum_{g=1}^{G_m} \exp(v_g) \right]$ and $x = \sum_{g=1}^{G_m} \exp(v_g - z) = \left[ \sum_{g=1}^{G_m} \exp(v_g) \right] \exp(-z) = V \exp(-z)$. I solve the above integral by substituting for $z = V - \ln x$, where $dz = -(1/x)dx$:

$$
E_x \left[ \max_{g \in G_m} v_g + \varepsilon_g \right] = \int_{-\infty}^{\infty} z \left( \sum_{g=1}^{G_m} \exp(v_g - z) \right) \exp \left( \sum_{g=1}^{G_m} - \exp(v_g - z) \right) dz
$$

$$
= \int_{-\infty}^{\infty} z \exp \left( \sum_{g=1}^{G_m} - \exp(v_g - z) \right) \left( \sum_{g=1}^{G_m} \exp(v_g - z) \right) dz
$$

$$
= \int_{-\infty}^{\infty} (V - \ln x) \exp (-x) x (-1/x) dx
$$

$$
= \int_{0}^{\infty} (V - \ln x) \exp (-x) dx
$$

$$
= V
$$

Since we have defined $\nu_{img} = (\gamma_m(Z)\beta_{mg} - \ln p_{mg})/\mu_m(Z)$ and $V = \ln \left[ \sum_{g=1}^{G_m} \exp(v_g) \right]$, we can use the above result to solve for the expectation in equation (A.11):

$$
E_x \left[ \ln \left( \max_{g \in G_m} \hat{p}_{img} \right) \right] = \mu_m(Z) \ln \left[ \sum_{g \in G_m} \exp((\gamma_m(Z)\beta_{mg} - \ln p_{mg})/\mu_m(Z)) \right]
$$

$$
= \mu_m(Z) \ln \left[ \sum_{g \in G_m} \left( \exp(\gamma_m(Z)\beta_{mg}) \right) \frac{1}{p_{mg}} \right] \ln (\mu_m(Z))
$$

$$
= \ln \left[ \sum_{g \in G_m} \left( \exp(\gamma_m(Z)\beta_{mg}) \right) \frac{1}{p_{mg}} \right] \mu_m(Z)
$$

(A.13)

Plugging this result back into equation (A.10) yields the expected relative module expenditure share for consumer $i$ in terms of product prices and model parameters:

$$
E_x \left[ \ln s_{im}(Z, P) - \ln s_{im}(Z, \overline{P}) \right]
$$

$$
= (\sigma(Z) - 1) E_x \left[ \ln \left( \max_{g \in G_m} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})}{p_{mg}} \right) \right]
$$

$$
- (\sigma(Z) - 1) E_x \left[ \ln \left( \max_{g \in G_m} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})}{p_{mg}} \right) \right]
$$

$$
= (\sigma(Z) - 1) \ln \left[ \sum_{g \in G_m} \left( \exp(\gamma_m(Z)\beta_{mg}) \right) \frac{1}{p_{mg}} \right] \mu_m(Z)
$$

$$
- (\sigma(Z) - 1) \ln \left[ \sum_{g \in G_m} \left( \exp(\gamma_m(Z)\beta_{mg}) \right) \frac{1}{p_{mg}} \right] \mu_m(Z)
$$

This function only varies by consumer through their outside good expenditure. All consumers with the same outside good expenditure and facing the same prices, $\overline{P}$, will have the same expected relative module expenditure.
The inclusive value, or price-adjusted quality index, for a household with income \( Y \) as a log linear function in the base product quality parameter, \( \beta \). Substituting this expression into equation (A.17), we can express the log of the inclusive value function as:

\[
V_m(Z, \mathbb{P}_m) = \left[ \sum_{g \in G_m} \left( \frac{\exp(\beta g_m (1 + \gamma_m \ln Z))}{p_{mg}} \right)^{-(\alpha_m^0 + \alpha_m^1 \ln Z)} \right]^{-\frac{1}{\alpha_m^0 + \alpha_m^1 \ln Z}} \tag{A.15}
\]

where I have substituted in the parametrizations for \( \gamma_m(Z) = (1 + \gamma_m \ln Z) \) and \( \mu_m(Z) = 1/ (\alpha_m^0 + \alpha_m^1 \ln Z) \). Equations (A.14) and (A.15) together define the expected relative module expenditure share of a set of households with income \( Y \) that face prices \( \mathbb{P}_m \) and \( \mathbb{P}_m \) in terms of parameters \( \alpha_m^0, \alpha_m^1 \), as well as \( \alpha_m, \gamma_m, \beta_mg \) for all \( g \in G_m \), and \( \alpha_m, \gamma_m, \beta_mg \) for all \( g \in G_m \).

### B.5 Extracting Second Stage Estimates \( \theta_2 \) From the Inclusive Value Function

Equation (25) relates the relative sample module expenditure shares to the relative log inclusive values for the modules:

\[
\ln \left( \frac{\mathbb{S}_{km}}{s_{km}} \right) = (\alpha^0 + \alpha^1 y_{kt}) \left[ \ln V_m(y_{kt}, \mathbb{P}_{kmt}) - \ln V_m(y_{kt}, \mathbb{P}_{kmt}) \right] + u_{kmmt} \tag{A.16}
\]

The inclusive value, or price-adjusted quality index, for a household with income \( y_{kt} \) facing a price vector \( \mathbb{P}_{kmt} \) in module \( m \) in market \( t \) is defined in equation 26 as:

\[
V_m(y_{kt}, \mathbb{P}_{kmt}) = \left[ \sum_{g \in G_m} \left( \frac{\mathbb{d}_{kmgt}}{p_{kmgt}} \right)^{-(\alpha_m^0 + \alpha_m^1 y_{kt})} \right]^{-\frac{1}{\alpha_m^0 + \alpha_m^1 y_{kt}}} \tag{A.17}
\]

where \( \mathbb{d}_{kmgt} = \exp(\beta g_m (1 + \gamma_m y_{kt})) \).

Note that the inclusive value is a function of the parameters estimated in both the first and second stage, i.e. \( \theta_1 \) and \( \theta_2 \). Specifically, each product quality, \( \beta_mg \), parameter is the sum of \((\beta_mg - \beta_{m\bar{g}})\), a component of the \( \theta_1 \) first-stage estimates, and \( \beta_{m\bar{g}} \), component of \( \theta_2 \) second-stage estimates.

Define \( \mathbb{d}_{kmgt} = \mathbb{d}_{kmgt} / \exp(\beta g_m (1 + \gamma_m y_{kt})) = \exp((\beta g_m - \beta_{m\bar{g}})(1 + \gamma_m y_{kt})) \) to reflect this decomposition. Substituting this expression into equation (A.17), we can express the log of the inclusive value function as a log linear function in the base product quality parameter, \( \beta_{m\bar{g}} \), to be estimated in the second stage:

\[
\ln V_m(y_{kt}, \mathbb{P}_{kmt}) = \ln \left[ \sum_{g \in G_m} \left( \frac{\mathbb{d}_{kmgt}}{p_{kmgt}} \right)^{-(\alpha_m^0 + \alpha_m^1 y_{kt})} \right]^{-\frac{1}{\alpha_m^0 + \alpha_m^1 y_{kt}}} + \beta_{m\bar{g}} (1 + \gamma_m y_{kt})
\]

Under the normalization that \( \beta_{m\bar{g}} = 0 \), and using the decomposition of the inclusive value function above, we can now rewrite equation (A.16) as:

\[
\ln \left( \frac{\mathbb{S}_{km}}{s_{km}} \right) = (\alpha^0 + \alpha^1 y_{kt}) \left[ \Delta V_{imn}(y_{kt}, \mathbb{P}_{kmt}, \mathbb{P}_{kmt}, \theta_1) + \beta_{m\bar{g}} (1 + \gamma_m y_{kt}) \right] + u_{kmmt},
\]
where $\Delta V_{1km}(y_{kt}, p_{kmt}, \theta_1) = \ln V_{1km}(y_{kt}, p_{kmt}, \theta_1) - \ln V_{1m}(y_{kt}, p_{kmt}, \theta_1)$ and 

$$V_{1m}(y_{kt}, p_{kmt}, \theta_1) = \left[ \sum_{g \in G_m} \left( \frac{d_{kmgt}}{p_{kmgt}} \right)^{-\epsilon_m^2 - \alpha_m^2} \right]^{-\epsilon_m^2 - \alpha_m^2 + \ln y_{kt}}. $$

C Connection to Nested CES Utility Function

In Section (1) of the paper, I model consumer demand assuming that a consumer $i$’s utility from grocery consumption, conditional on their outside good expenditure $Z$, is a CES aggregate over consumer-specific module-level utilities that are, in turn, additive in product-level log-logit utilities. This utility function is presented in equations (1), (2), and (3) and can be summarized as:

$$U_{iG}(Q, Z) = \left\{ \sum_{m \in M} u_{im}(Q_m, Z) \right\}^{\sigma(Z) / (\sigma(Z) - 1)} = \left\{ \sum_{m \in M} \left( \sum_{g \in G_m} u_{img}(Q_m, Z) \right) \right\}^{\sigma(Z) / (\sigma(Z) - 1)} = \left\{ \sum_{m \in M} \left( \sum_{g \in G_m} q_{mg} \exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\epsilon_{img}) \right) \right\}^{\sigma(Z) / (\sigma(Z) - 1)} = \left\{ \sum_{m \in M} \left( \sum_{g \in G_m} q_{mg} \exp(\gamma_m(Z)\beta_{mg} + \frac{\epsilon_{img}}{\sigma_m(Z) - 1}) \right) \right\}^{\sigma(Z) / (\sigma(Z) - 1)} (A.18)$$

where $q_{mg}$ is the consumption quantity of each product $g$ in module $m$; $\beta_{mg}$ is the quality of product $g$ in module $m$; $\epsilon_{img}$ is the idiosyncratic utility of consumer $i$ from product $g$ in module $m$; $\gamma_m(Z)$ and $\mu_m(Z) = \frac{1}{\sigma_m(Z) - 1} > 0$ are weights that govern the extent to which consumers with outside good expenditure $Z$ care about product quality and their idiosyncratic utility draws; and $\sigma_m(Z)$ is the elasticity of substitution between products in the same module $m$ and $\sigma(Z) > 1$ is the elasticity of substitution between products in different modules for a consumer with outside good expenditure $Z$.

Consider the utility of the representative agent for consumers with outside good expenditure $Z$. This agent’s utility function from grocery consumption is defined in equation (16) in Section (5.1) as follows:

$$U_{iG}^{CES}(Q, Z) = \left\{ \sum_{m \in M} \left[ \sum_{g \in G_m} [q_{mg} \exp(\beta_{mg} \gamma_m(Z))] \right]^{\sigma_m(Z) / (\sigma_m(Z) - 1)} \right\}^{\sigma(Z) / (\sigma(Z) - 1)} (A.19)$$

where $q_{mg}$, $\beta_{mg}$, $\gamma_m(Z)$, $\sigma_m(Z)$, and $\sigma(Z) > 1$ take the same definition as in equation (A.18) above.

Suppose that this representative consumer with the nested-CES utility function $U_{iG}^{CES}(Q, Z)$ defined in equation (A.19) faces the same prices $P$ and has the same outside good expenditure $Z$ as a group of “idiosyncratic” consumers with the CES-nested log-logit utility $U_{iG}(Q, Z)$ defined in equation (A.18). A simple extension of
Anderson et al. (1987) shows that the representative consumer and the group of “idiosyncratic” consumers will allocate expenditures across products within modules and across modules identically.

First consider the within-module expenditure allocations. Denote the share of module \( m \) expenditures that the representative consumer allocates to product \( g \) as \( s_{mg|m}^{CES}(Z) \) and the share of total grocery expenditures the representative consumer allocates to module \( m \) as \( s_m^{CES}(Z) \). This share is equal to

\[
s_{mg|m}^{CES}(Z) = \left[ \frac{Pmg}{\exp(\beta mg \gamma_m(Z)) / P_m^{CES}(Z, \bar{P}_m)} \right]^{1-\sigma_m(Z)}
\]

where \( P_m^{CES}(Z, \bar{P}_m) \) is a module-level CES price index. The relative log share that the representative consumer optimally allocates to product \( g \) in module \( m \) relative to some other product \( \bar{g} \) in the same module is, therefore,

\[
\ln s_{mg|m}^{CES}(Z) - \ln s_{mg|m}^{CES}(Z) = (1 - \sigma_m(Z)) \left[ (\ln p_{mg} - \ln p_{mg}) - (\beta mg - \beta mg) \gamma_m(Z) \right]
\]

The expected relative module expenditure share of a group of “idiosyncratic” consumers with outside good expenditure \( Z \) facing the same sets of prices \( p_{mg} \) and \( p_{mg} \) is derived in Appendix (B.3) as:

\[
\mathbb{E}_z \left[ \ln(s_{mg|m}(Z, \bar{P}_m)) - \ln(s_{mg|m}(Z, \bar{P}_m)) \right] = (\sigma_m(Z) - 1) \left[ (\beta mg - \beta mg) \gamma_m(Z) - (\ln p_{mg} - \ln p_{mg}) \right]
\]

where I have substituted \( \sigma_m(Z) \) and \( \gamma_m(Z) \) for their log-linear parametric forms \((1 + \alpha_0 + \alpha_1 \ln Z)\) and \((1 + \gamma_m \ln Z)\), respectively. We can multiply both terms of the right-hand side of (A.21) to show that it is equivalent to the right-hand side of equation (A.20):

\[
\mathbb{E}_z \left[ \ln(s_{mg|m}(Z, \bar{P}_m)) - \ln(s_{mg|m}(Z, \bar{P}_m)) \right] = (\sigma_m(Z) - 1) \left[ (\beta mg - \beta mg) \gamma_m(Z) - (\ln p_{mg} - \ln p_{mg}) \right] = (1 - \sigma_m(Z)) \left[ (\ln p_{mg} - \ln p_{mg}) - (\beta mg - \beta mg) \gamma_m(Z) \right] = \ln s_{mg|m}^{CES}(Z) - \ln s_{mg|m}^{CES}(Z)
\]

whereby showing that the representative consumer allocates expenditures across products in the same module identically to a group of the “idiosyncratic” consumers.

Now consider the between-module expenditure allocations. Denote the share of total grocery expenditures the representative consumer allocates to module \( m \) as \( s_m^{CES}(Z) \). The relative log share that the representative consumer optimally allocates to module \( m \) relative to some other module \( \bar{m} \) is

\[
\ln s_m^{CES}(Z) - \ln s_m^{CES}(Z) = (1 - \sigma(Z)) \left[ \ln \left( P_m^{CES}(Z, \bar{P}_m) \right) - \ln \left( P_m^{CES}(Z, \bar{P}_m) \right) \right]
\]

where \( P_m^{CES}(Z, \bar{P}_m) \) is a module-level CES price index defined as:

\[
P_m^{CES}(Z, \bar{P}_m) = \left[ \sum_{g \in G_m} \left( \frac{Pmg}{\exp(\beta mg \gamma_m(Z))} \right)^{(1-\sigma_m(Z))} \right]^{-1/(1-\sigma_m(Z))}
\]

The expected relative module expenditure share of a group of “idiosyncratic” consumers with outside good expenditure \( Z \) facing the same sets of prices \( \bar{P}_m \) and \( P_m \) faced by the representative consumer is derived in Appendix
(B.4) as:

\[ E[\ln s_{im}(Z, P) - \ln s_{im}(Z, P)] = (\sigma(Z) - 1) [\ln V_m(Z, P_m) - \ln V_m(Z, P_m)] \tag{A.24} \]

where \( V_m(Z, P_m) \) is a CES-style index over price-adjusted product qualities:

\[ V_m(Z, P_m) = \left[ \sum_{g \in G_m} \left( \frac{\exp(\beta_m \gamma_m(Z))}{p_{mg}} \right)^{\sigma_m(Z) - 1} \right]^{\frac{1}{\sigma_m(Z) - 1}} \tag{A.25} \]

To see that the right-hand sides of equations (A.22) and (A.24) are identical first note that we can re-write the equation (A.24) as

\[ E[\ln s_{im}(Z, P) - \ln s_{im}(Z, P)] = (1 - \sigma(Z)) \left[ -\ln V_m(Z, P_m) + \ln V_m(Z, P_m) \right] \]

\[ = (1 - \sigma(Z)) \left[ \ln \left( V_m(Z, P_m) \right) - \ln \left( V_m(Z, P_m) \right)^{-1} \right] \]

In fact, the right-hand sides of equations (A.22) and (A.24) will be identical as long as the quality-adjusted price levels defined in equation (A.23) are equal to the inverse of the price-adjusted quality levels defined in equation (A.25), i.e., \( P_{CES}^m(Z, P_m) = \left[ V_m(Z, P_m) \right]^{-1} \). We can see this is the case below:

\[ P_{CES}^m(Z, P_m) = \left[ \sum_{g \in G_m} \left( \frac{p_{mg}}{\exp(\beta_m \gamma_m(Z))} \right)^{1 - \sigma_m(Z)} \right]^{\frac{1}{1 - \sigma_m(Z)}} \]

\[ = \left[ \sum_{g \in G_m} \left( \frac{\exp(\beta_m \gamma_m(Z))}{p_{mg}} \right)^{\sigma_m(Z) - 1} \right]^{\frac{1}{1 - \sigma_m(Z)}} \]

\[ = \left[ \sum_{g \in G_m} \left( \frac{\exp(\beta_m \gamma_m(Z))}{p_{mg}} \right)^{\sigma_m(Z) - 1} \right]^{-1} \]

\[ = \left[ V_m(Z, P_m) \right]^{-1} \]

The representative consumer therefore allocates expenditures across modules in identical proportions to a group of the “idiosyncratic” consumers.

The algebra above has shown that the CES-nested log-logit utility function yields identical relative expenditure share equations, both across and within modules, to the nested-CES utility function assumed for the representative agent. In particular, note that the model parameters play identical roles in the nested-CES and CES-nested log-logit expenditure share equations, so the parameter estimates identified using moments based on these equations can be used as direct inputs into the nested-CES price indexes that form the basis for the main results presented above.
D Procedure for Obtaining Standard Errors of Upper-Level
Demand Parameters

I estimate the parameters of the model sequentially. Recall that the full set of demand parameters, \( \theta \), are partitioned into \( M \) sets of lower-level module-specific parameters, \( \theta_{1m} \), for each module \( m \), that are identified using module-specific sub-samples of the data and a single set of parameters, \( \theta_2 \), whose identification requires data from all modules. Newey and McFadden (1994) show how to obtain a consistent covariance matrix for estimates that are obtained sequentially and Murphy and Topel (1985) describe the assumptions under which this method can be extended to the case in which the first-step estimates are obtained from different models estimated using sub-samples of the data. In this Appendix, I outline how I apply these methods to calculate the covariance matrix of the upper-level demand parameters.

D.1 Step 1: Parallel Estimation of \( \theta_1 \)

The first step in my estimation is to obtain estimates for \( \theta_1 = \left\{ \alpha_0^m, \alpha_1^m, \gamma_m, \left\{ \beta_{mg} - \beta_m \bar{g}_m \right\} \right\}_{m=1,...,M} \), where \( \tilde{\beta}_{mg} \) denotes \( \beta_{mg} - \beta_m \bar{g}_m \). I obtain \( \hat{\theta}_1 \) using a two-stage GMM procedure based on the following exogeneity restriction:

\[
E[f(X; \theta_1)] = 0 \tag{A.26}
\]

where \( f(X; \theta_1) = Z_1(\mathbf{X})'\nu(X; \theta_1), Z_1(\mathbf{X}) \) is a stacked vector of \( L_{1m} \) module-specific instruments, \( Z_{1m}(\mathbf{X}) \), for each module \( m \). \( \nu(X; \theta_1) \) is the error in the relative within-module expenditure share equation. For income group, \( k \), market \( t \), and product \( g \) in module \( m \), this error is defined as:

\[
\nu_{gkt}(\mathbf{X}_m; \theta_{1m}) = \ln \left( \frac{s_{gkt}}{s_{gkt}} \right) - (\alpha_0^m + \alpha_1^m y_{kt}) \left[ \tilde{\beta}_{mg} (1 + \gamma_m y_{kt}) - \ln \left( \frac{p_{gkt}}{p_{\bar{g}_m k t}} \right) \right] \tag{A.27}
\]

The fact that these errors depend only on module-specific data, \( \mathbf{X}_m \), and parameters, \( \theta_{1m} \), enables me to partition A.26 into module-specific auxiliary moments:

\[
E[f(\mathbf{X}_m; \theta_{1m})] = 0
\]

for \( f(\mathbf{X}_m; \theta_{1m}) = Z_{1m}(\mathbf{X}_m)'\nu(\mathbf{X}_m; \theta_{1m}) \).

This partition allows me to estimate the \( K_{1m} \) parameters, \( \theta_{1m} = \left\{ \alpha_0^m, \alpha_1^m, \gamma_m, \left\{ \tilde{\beta}_{mg} \right\} \right\}_{g \in G_m, g \neq \bar{g}_m} \), for each module \( m \) in separate but parallel minimization procedures. Consistent estimates are obtained by minimizing module-specific GMM objective functions as follows:

\[
\hat{\theta}_{1m} = \arg \min_{\theta_{1m}} \sum_{k,t} N_{kt} \sum_{g \in G_{mkt}} f_{gkt}(\mathbf{X}_m; \theta_{1m})
\]

where \( f(\mathbf{X}_m; \theta_{1m}) = \sum_{k,t} \sum_{g \in G_{mkt}} f_{gkt}(\mathbf{X}_m; \theta_{1m}) \) is the sample analog of \( E[f(\mathbf{X}_m; \theta_{1m})] \); \( f_{gkt}(\mathbf{X}_m; \theta_{1m}) = Z_{1gkt}(\mathbf{X}_m)'\nu_{gkt}(\mathbf{X}_m; \theta_{1m}) \); \( G_{mkt} \) is the set of \( N_{mkt} \) module \( m \) non-base (i.e., \( g \neq \bar{g}_m \)) products purchased by income-group \( k \) households in market \( t \); and \( Z_{1gkt} \) is the \( 1 \times L_{1m} \) \((L_{1m} \leq K_{1m})\) vector of instruments for a prod-
uct $g$-income group $k$-market $t$ observation. $W_{1m} = \left[ \frac{1}{N_{mkt}} \sum_{k,t} \sum_{g \in G_{mkt}} f_{gkt}(X_m; \hat{\theta}_{1m}) f_{gkt}(X_m; \hat{\theta}_{1m})' \right]^{-1}$ is the efficient weighting matrix, calculated using consistent first-stage estimates of $\hat{\theta}_{1m}$:

$$\hat{\theta}_{1m} = \arg \min_{\hat{\theta}_{1m}} \hat{J}(X_m; \hat{\theta}_{1m})' W_{1m} \hat{J}(X_m; \hat{\theta}_{1m})$$

for $W_{1m} = \left[ \frac{1}{N_{mkt}} \sum_{k,t} \sum_{g \in G_{mkt}} Z_{1gkt}Z_{1gkt}' \right]^{-1}$.

Assuming that the random components of the $M$ module-specific auxiliary models are independent, the variance-covariance matrix of $\hat{\theta}_1$, $\Omega_1$, can be written as:

$$\Omega_\theta = \begin{bmatrix} \Omega_{\hat{\theta}_{11}} & 0 \\ 0 & \ddots & \vdots \\ 0 & \ddots & \Omega_{\hat{\theta}_{1M}} \end{bmatrix}$$

where $\Omega_{\hat{\theta}_{1m}}$ is the variance-covariance matrix of $\hat{\theta}_{1m}$ for each $m = 1, \ldots, M$. The consistent estimator for each of these sub-matrices is:

$$\hat{\Omega}_{\hat{\theta}_{1m}} = \left( \hat{F}_{\hat{\theta}_{1m}} \hat{V}_{ff}^{-1} \hat{F}_{\hat{\theta}_{1m}}' \right)^{-1}$$

where

$$\hat{F}_{\hat{\theta}_{1m}} = \frac{1}{N_{mkt}} \sum_{k,t} \sum_{g \in G_{mkt}} \nabla_{\hat{\theta}_{1m}} f_{gkt}(X_m; \hat{\theta}_{1m}) (K_1 \times L_1)$$

and

$$\hat{V}_{ff} = W_1 = \frac{1}{N_{mkt}} \sum_{k,t} \sum_{g \in G_{mkt}} f_{gkt}(X_m; \hat{\theta}_{1m}) f_{gkt}(X_m; \hat{\theta}_{1m})' (L_1 \times L_1).$$

**D.2 Step 2: Sequential Estimation of $\theta_2$**

In the second step of the sequential estimation procedure, I estimate $\theta_2 = \{ \alpha^0, \alpha^1, \{ \beta_{m\bar{m}} \}_{m = 1, \ldots, M, m \neq \bar{m}} \}$. These $K_2 = 1 + M$ parameters are identified by the following exogeneity restriction:

$$G = \mathbb{E}[h(X; \hat{\theta}_1, \hat{\theta}_2)] = 0 \quad (A.28)$$

where $h(X; \hat{\theta}_1, \hat{\theta}_2) = Z_2(X) \cdot u(X; \hat{\theta}_1, \hat{\theta}_2)$. $Z_2(X)$ is a set of $L_2$ instruments ($L_2 \geq K_2$) and $u(X; \hat{\theta}_1, \hat{\theta}_2)$ is the error in the relative across-module expenditure share equation. For income group, $k$, market $t$, and module $m$, this error is defined as:

$$u_{kmt}(X; \hat{\theta}_1, \hat{\theta}_2) = \ln \left( \frac{Z_{mkt}}{\bar{Z}_{mkt}} \right) - (\alpha^0 + \alpha^1 y_{kt}) [\Delta V_{1m\bar{m}}(y_{kt}, P_{mkt}, \bar{P}_{\bar{m}kt}, \hat{\theta}_1) + \beta_{m\bar{m}} (1 + \gamma_m y_{kt})],$$

64
where
\[ \Delta V_{1m}(y_{kt}, \mathbb{P}_{mkt}, \mathbb{P}_{mkt}; \theta_1) \triangleq \ln V_{1m}(y_{kt}, \mathbb{P}_{mkt}; \theta_1) - \ln V_{1m}(y_{kt}, \mathbb{P}_{mkt}; \theta_1) , \]
and
\[ V_{1m}(y_{kt}, \mathbb{P}_{mkt}; \theta_1) = \sum_{g \in G_m} \left( \exp \left( \tilde{\beta}_{mg} (1 + \gamma_m y_{kt}) \right) \right) ^ {-\left( \alpha_m^1 + \alpha_m^0 y_{kt} \right)} \]

The first stage \( \hat{\theta}_1 \) estimates are inputs into the sample moment condition used to estimate the \( 1 \times K_2 \) vector of \( \theta_2 \) parameters, denoted \( \hat{\theta}_2 \). These upper-level parameters are estimated using two-step GMM:
\[ \hat{\theta}_2 = \arg \min_{\theta_2} h(\mathbf{X}; \hat{\theta}_1, \theta_2) \hat{W}_2 h(\mathbf{X}; \hat{\theta}_1, \theta_2) \]
where \( h(\mathbf{X}; \hat{\theta}_1, \theta_2) = \sum_{k,t} \sum_{m \in M_{kt}} h_{mkt}(\mathbf{X}; \hat{\theta}_1, \theta_2) \) is the sample analog of \( E[h(\mathbf{X}; \hat{\theta}_1, \theta_2)] \); \( h_{mkt}(\mathbf{X}; \hat{\theta}_1, \theta_2) = z_{2mkt}^u mkt(\mathbf{X}; \hat{\theta}_1, \theta_2); mkt \) is the set of \( N_{kt} \) non-base modules (i.e., \( m \neq \bar{m} \)) purchased by income-group \( k \) households in market \( t \); and \( \mathbf{Z}_{2mkt} \) is the \( 1 \times L_2 \) vector of instruments for a module \( m \)-income group \( k \)-market \( t \) observation. \( \hat{W}_2 = \left( \sum_{k,t} \sum_{m \in M_{kt}} h_{mkt}(\mathbf{X}; \hat{\theta}_1, \hat{\theta}_2) h_{mkt}(\mathbf{X}; \hat{\theta}_1, \hat{\theta}_2)' \right)^{-1} \) is the optimal weighting matrix, where \( \hat{\theta}_2 \) are consistent first-stage estimates of \( \theta_2 \) that minimize a GMM objective function as follows:
\[ \hat{\theta}_2 = \arg \min_{\theta_2} h(\mathbf{X}; \hat{\theta}_1, \theta_2) \hat{W}_2 h(\mathbf{X}; \hat{\theta}_1, \theta_2) \]
for \( \hat{W}_2 = \left( \sum_{k,t} \sum_{m \in M_{kt}} \mathbf{Z}_{2mkt} \mathbf{Z}_{2mkt}' \right)^{-1} \).

The naive variance-covariance matrix of the \( \hat{\theta}_2 \) estimates that does not account for the measurement error from the use of the first stage estimates, treating \( \theta_1 \) as known, is defined as:
\[ \hat{\Omega}_{\theta_2} = \left( \hat{K}_{\theta_2} \hat{V}_{\theta_1}^{-1} \hat{K}_{\theta_2}' \right)^{-1} \]
where
\[ \hat{K}_{\theta_2} = \frac{1}{N_{kt}} \sum_{k,t} \sum_{m \in M_{kt}} \nabla_{\theta_2} h_{mkt}(\mathbf{X}; \hat{\theta}_1, \hat{\theta}_2) (K_2 \times L_2) \]
and
\[ \hat{V}_{\theta_1} = \hat{W}_2 = \frac{1}{N_{kt}} \sum_{k,t} \sum_{m \in M_{kt}} h_{mkt}(\mathbf{X}; \hat{\theta}_1, \hat{\theta}_2) h_{mkt}(\mathbf{X}; \hat{\theta}_1, \hat{\theta}_2)' (L_2 \times L_2) . \]

In order to account for the measurement error from the use of first state estimates, we need to treat \( \theta_1 \) as
unknown, calculating the variance-covariance of the full vector of \( \hat{\theta} \) estimates:

\[
\Omega_{\theta} = \begin{bmatrix}
\Omega_{\theta_1} & \Omega_{\theta_1\theta_2} \\
\Omega_{\theta_1\theta_2} & \Omega_{\theta_2}
\end{bmatrix} = (C_\theta V_{cc}^{-1} C_\theta')^{-1}
\]

where:

\[
C_\theta = \begin{bmatrix}
F_{\theta_1} & 0 \\
H_{\theta_1} & H_{\theta_2}
\end{bmatrix}
\quad \text{and} \quad
V_{cc} = \begin{bmatrix}
V_{ff} & V_{fh} \\
V_{fh} & V_{hh}
\end{bmatrix}
\]

The correct covariance matrix for the second stage estimates is the lower right-hand block of this full covariance matrix, \( \Omega_{\theta_2} \). I obtain it by estimating the full covariance matrix, \( \hat{\Omega}_{\theta_2} \), where \( \hat{\Omega}_{\theta_1}, \hat{\Omega}_{\theta_2}, \hat{H}_{\theta_2}, \text{ and } \hat{F}_{\theta_1} \) are as defined above:

\[
\hat{H}_{\theta_1} = \frac{1}{\sum_{k,t} N_{kt} \sum_{k,t} \sum_{m \in M_{kt}} \nabla_{\theta_1} h_{mk}(X; \hat{\theta}_1, \hat{\theta}_2) (K_1 \times L_2);
\]

and

\[
\hat{V}_{fg} = \hat{V}_{gf}' = \frac{1}{\sum_{k,t} N_{kt} \sum_{k,t} \sum_{m \in M_{kt}} h_{mk}(X; \hat{\theta}_1, \hat{\theta}_2)} \frac{1}{\sum_{g \in G_{mkt}} f_{gkt}(X; \hat{\theta}_{1m})}' (L_2 \times L_1).
\]

\[53\] Newey (1984) shows that, when \( L_1 = K_1 \) and \( L_2 = K_2 \), the asymptotic covariance matrix \( \Omega_{\theta_2} \) of the second step estimator \( \hat{\theta}_2 \) is given by:

\[
\hat{\Omega}_{\theta_2} = \hat{\Omega}_{\theta_2} + \hat{H}_{\theta_2}^{-1} \hat{H}_{\theta_1} \hat{\Omega}_{\theta_1} \hat{H}_{\theta_1}' \hat{H}_{\theta_2}^{-1} \left( \hat{H}_{\theta_1} \hat{F}_{\theta_1}^{-1} \hat{V}_{fh} + \hat{V}_{fh} (\hat{F}_{\theta_1})' (\hat{H}_{\theta_1})^{-1} \right)
\]

where \( \hat{\Omega}_{\theta_1}, \hat{\Omega}_{\theta_2}, \hat{H}_{\theta_2}, \text{ and } \hat{F}_{\theta_1} \) are as defined above and \( \hat{H}_{\theta_1} = \sum_{k,t} N_{kt} \sum_{k,t} \sum_{m \in M_{kt}} \nabla_{\theta_1} h_{mk}(X; \hat{\theta}_1, \hat{\theta}_2) \). This equation cannot be applied directly to estimate \( \Omega_{\theta_2} \) here since both models estimated here are over-identified, such that \( L_1 > K_1 \) and \( L_2 > K_2 \) (and neither \( \hat{F}_{\theta_1} \) or \( \hat{H}_{\theta_2} \) are invertible).
### Figure A.2: Distribution of $\gamma$ Parameter Estimates Across Modules

3. Non-Homothetic in Quality

4. Non-Homothetic in Price and Quality

Income-Quality Weight Gradient (gamma)

### Figure A.3: Distribution of $\alpha_1$ Parameter Estimates Across Modules

2. Non-Homothetic in Price

4. Non-Homothetic in Price and Quality

Coefficient on $-\text{Ln(Relative Price)} \times \text{Ln(Income)}$ (alpha1)
Table A.1: Sample Size, Population, and Income by Market

<table>
<thead>
<tr>
<th>Market ID</th>
<th>Market Name</th>
<th>Sample Household Count in 2005</th>
<th>Population</th>
<th>Per Capita Income</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Des Moines</td>
<td>143</td>
<td>45,58</td>
<td>456,022</td>
<td>37,650</td>
</tr>
<tr>
<td>2</td>
<td>Little Rock</td>
<td>348</td>
<td>185,59</td>
<td>583,845</td>
<td>33,289</td>
</tr>
<tr>
<td>3</td>
<td>Omaha</td>
<td>116</td>
<td>41,50</td>
<td>716,998</td>
<td>37,869</td>
</tr>
<tr>
<td>4</td>
<td>Syracuse</td>
<td>164</td>
<td>73,57</td>
<td>732,117</td>
<td>31,445</td>
</tr>
<tr>
<td>5</td>
<td>Albany</td>
<td>135</td>
<td>47,48</td>
<td>875,583</td>
<td>24,811</td>
</tr>
<tr>
<td>6</td>
<td>Birmingham</td>
<td>565</td>
<td>179,230</td>
<td>921,106</td>
<td>35,448</td>
</tr>
<tr>
<td>7</td>
<td>Richmond</td>
<td>192</td>
<td>51,77</td>
<td>996,512</td>
<td>37,082</td>
</tr>
<tr>
<td>8</td>
<td>Louisville</td>
<td>449</td>
<td>150,192</td>
<td>1,025,598</td>
<td>34,162</td>
</tr>
<tr>
<td>9</td>
<td>Grand Rapids</td>
<td>198</td>
<td>84,66</td>
<td>1,088,514</td>
<td>31,966</td>
</tr>
<tr>
<td>10</td>
<td>Jacksonville</td>
<td>128</td>
<td>39,45</td>
<td>1,100,491</td>
<td>35,439</td>
</tr>
<tr>
<td>11</td>
<td>Memphis</td>
<td>328</td>
<td>117,125</td>
<td>1,135,614</td>
<td>34,052</td>
</tr>
<tr>
<td>12</td>
<td>Raleigh-Durham</td>
<td>238</td>
<td>79,86</td>
<td>1,187,941</td>
<td>35,858</td>
</tr>
<tr>
<td>13</td>
<td>Nashville</td>
<td>397</td>
<td>154,156</td>
<td>1,231,311</td>
<td>36,056</td>
</tr>
<tr>
<td>14</td>
<td>Salt Lake City</td>
<td>245</td>
<td>92,103</td>
<td>1,333,914</td>
<td>33,469</td>
</tr>
<tr>
<td>15</td>
<td>Charlotte</td>
<td>1,100</td>
<td>392,403</td>
<td>1,499,293</td>
<td>36,580</td>
</tr>
<tr>
<td>16</td>
<td>Columbus</td>
<td>1,095</td>
<td>402,293</td>
<td>1,540,157</td>
<td>34,777</td>
</tr>
<tr>
<td>17</td>
<td>San Antonio</td>
<td>911</td>
<td>301,345</td>
<td>1,592,383</td>
<td>31,189</td>
</tr>
<tr>
<td>18</td>
<td>Indianapolis</td>
<td>243</td>
<td>86,92</td>
<td>1,607,486</td>
<td>36,160</td>
</tr>
<tr>
<td>19</td>
<td>Orlando</td>
<td>246</td>
<td>108,78</td>
<td>1,644,561</td>
<td>31,828</td>
</tr>
<tr>
<td>20</td>
<td>Milwaukee</td>
<td>169</td>
<td>57,64</td>
<td>1,689,572</td>
<td>37,361</td>
</tr>
<tr>
<td>21</td>
<td>Hartford-New Haven</td>
<td>193</td>
<td>62,69</td>
<td>1,725,259</td>
<td>41,705</td>
</tr>
<tr>
<td>22</td>
<td>Kansas City</td>
<td>219</td>
<td>83,80</td>
<td>1,776,062</td>
<td>35,839</td>
</tr>
<tr>
<td>23</td>
<td>Sacramento</td>
<td>1,044</td>
<td>297,395</td>
<td>1,796,857</td>
<td>35,318</td>
</tr>
<tr>
<td>24</td>
<td>New Orleans-Mobile</td>
<td>406</td>
<td>168,151</td>
<td>1,877,984</td>
<td>21,446</td>
</tr>
<tr>
<td>25</td>
<td>Oklahoma City-Tulsa</td>
<td>520</td>
<td>226,177</td>
<td>1,886,581</td>
<td>34,068</td>
</tr>
<tr>
<td>26</td>
<td>Cincinnati</td>
<td>316</td>
<td>97,112</td>
<td>1,979,202</td>
<td>35,326</td>
</tr>
<tr>
<td>27</td>
<td>Portland, Or</td>
<td>259</td>
<td>115,87</td>
<td>2,265,223</td>
<td>34,921</td>
</tr>
<tr>
<td>28</td>
<td>Buffalo-Rochester</td>
<td>1,121</td>
<td>476,384</td>
<td>2,268,312</td>
<td>33,020</td>
</tr>
<tr>
<td>29</td>
<td>Pittsburgh</td>
<td>300</td>
<td>150,100</td>
<td>2,358,695</td>
<td>36,159</td>
</tr>
<tr>
<td>30</td>
<td>Tampa</td>
<td>1,162</td>
<td>468,420</td>
<td>2,395,997</td>
<td>33,678</td>
</tr>
<tr>
<td>31</td>
<td>Denver</td>
<td>1,118</td>
<td>313,414</td>
<td>2,581,506</td>
<td>42,476</td>
</tr>
<tr>
<td>32</td>
<td>St. Louis</td>
<td>1,174</td>
<td>394,331</td>
<td>2,603,607</td>
<td>35,991</td>
</tr>
<tr>
<td>33</td>
<td>San Diego</td>
<td>138</td>
<td>47,38</td>
<td>2,813,833</td>
<td>40,383</td>
</tr>
<tr>
<td>34</td>
<td>Cleveland</td>
<td>360</td>
<td>137,93</td>
<td>2,945,831</td>
<td>35,555</td>
</tr>
<tr>
<td>35</td>
<td>Minneapolis</td>
<td>1,229</td>
<td>367,437</td>
<td>2,968,806</td>
<td>42,457</td>
</tr>
<tr>
<td>36</td>
<td>Phoenix</td>
<td>1,204</td>
<td>478,306</td>
<td>3,251,876</td>
<td>32,660</td>
</tr>
<tr>
<td>37</td>
<td>Seattle</td>
<td>1,139</td>
<td>302,406</td>
<td>3,554,760</td>
<td>42,356</td>
</tr>
<tr>
<td>38</td>
<td>Miami</td>
<td>1,069</td>
<td>315,359</td>
<td>3,876,380</td>
<td>38,342</td>
</tr>
<tr>
<td>39</td>
<td>Atlanta</td>
<td>1,029</td>
<td>253,383</td>
<td>4,112,198</td>
<td>35,262</td>
</tr>
<tr>
<td>40</td>
<td>Houston</td>
<td>1,065</td>
<td>295,416</td>
<td>4,669,571</td>
<td>40,734</td>
</tr>
<tr>
<td>41</td>
<td>Dallas</td>
<td>891</td>
<td>237,351</td>
<td>5,221,801</td>
<td>38,089</td>
</tr>
<tr>
<td>42</td>
<td>Detroit</td>
<td>991</td>
<td>261,358</td>
<td>5,456,428</td>
<td>37,204</td>
</tr>
<tr>
<td>43</td>
<td>Boston</td>
<td>1,271</td>
<td>348,484</td>
<td>5,819,100</td>
<td>47,491</td>
</tr>
<tr>
<td>44</td>
<td>Philadelphia</td>
<td>1,169</td>
<td>327,420</td>
<td>6,188,463</td>
<td>40,948</td>
</tr>
<tr>
<td>45</td>
<td>San Francisco</td>
<td>991</td>
<td>175,507</td>
<td>7,039,362</td>
<td>54,191</td>
</tr>
<tr>
<td>46</td>
<td>Washington, DC-Baltimore</td>
<td>1,106</td>
<td>238,520</td>
<td>7,608,070</td>
<td>46,725</td>
</tr>
<tr>
<td>47</td>
<td>Chicago</td>
<td>1,056</td>
<td>279,419</td>
<td>9,157,540</td>
<td>39,454</td>
</tr>
<tr>
<td>48</td>
<td>Los Angeles</td>
<td>1,143</td>
<td>289,462</td>
<td>16,373,645</td>
<td>37,441</td>
</tr>
<tr>
<td>49</td>
<td>New York</td>
<td>1,460</td>
<td>340,669</td>
<td>21,199,865</td>
<td>46,221</td>
</tr>
</tbody>
</table>
Table A.2: Bilateral Model Comparisons
Sales Share of Modules where GMM-BIC(M)<GMM-BIC(N)

<table>
<thead>
<tr>
<th>Model</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Homothetic</td>
<td>-</td>
<td>0.18</td>
<td>0.28</td>
<td>0.32</td>
</tr>
<tr>
<td>2. Non-Homothetic in Price</td>
<td>0.82</td>
<td>-</td>
<td>0.38</td>
<td>0.40</td>
</tr>
<tr>
<td>3. Non-Homothetic in Quality</td>
<td>0.72</td>
<td>0.62</td>
<td>-</td>
<td>0.68</td>
</tr>
<tr>
<td>4. Non-Homothetic in Quality and Price</td>
<td>0.68</td>
<td>0.60</td>
<td>0.32</td>
<td>-</td>
</tr>
</tbody>
</table>

Table A.3: City-Income Specific Price Index Regressions
Dependent Variable: \( \text{Ln(Price Index for Representative Consumer } k \text{ in City } c) \)

<table>
<thead>
<tr>
<th>Model allowing for NH in:</th>
<th>Quality</th>
<th>Price and Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[1]</td>
<td>[2]</td>
</tr>
<tr>
<td>Ln(Per Capita Income(_c))</td>
<td>( \beta_1 ) 2.412***</td>
<td>2.290*</td>
</tr>
<tr>
<td></td>
<td>[0.996]</td>
<td>[1.22]</td>
</tr>
<tr>
<td>Ln(Per Capita Income(_c))</td>
<td>( \beta_2 ) -0.217**</td>
<td>-0.201*</td>
</tr>
<tr>
<td>( *\text{Ln(Household Income}_k) )</td>
<td>[0.0915]</td>
<td>[0.112]</td>
</tr>
<tr>
<td>Ln(Population(_c))</td>
<td>( \beta_3 ) -</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>[0.232]</td>
<td>-</td>
</tr>
<tr>
<td>Ln(Population(_c))</td>
<td>( \beta_4 ) -</td>
<td>-0.005</td>
</tr>
<tr>
<td>( *\text{Ln(Household Income}_k) )</td>
<td>-</td>
<td>[0.021]</td>
</tr>
</tbody>
</table>

Household Income Fixed Effects | Yes | Yes | Yes | Yes |
Observations | 230 | 230 | 230 | 230 |
R-Squared | 0.03 | 0.036 | 0.04 | 0.035 |

*** p<0.01, ** p<0.05, * p<0.1; Standard errors in brackets.
### Table A.5: City-Specific Price Indexes Calculated Using Homothetic and Non-Homothetic Models

<table>
<thead>
<tr>
<th>Market</th>
<th>Homothetic</th>
<th>Non-Homothetic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$16,896</td>
<td>$26,715</td>
</tr>
<tr>
<td></td>
<td>$35,715</td>
<td>$41,526</td>
</tr>
<tr>
<td></td>
<td>$53,103</td>
<td>$60,442</td>
</tr>
<tr>
<td></td>
<td>$64,805</td>
<td>$82,576</td>
</tr>
<tr>
<td></td>
<td>$93,411</td>
<td>$146,566</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>0.83</td>
<td>0.77</td>
</tr>
<tr>
<td>Washington, DC-Baltimore</td>
<td>0.83</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>0.87</td>
<td>0.88</td>
</tr>
<tr>
<td>Atlanta</td>
<td>0.93</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.77</td>
</tr>
<tr>
<td>Detroit</td>
<td>0.95</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.89</td>
</tr>
<tr>
<td>Columbus</td>
<td>0.98</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td>Tampa</td>
<td>0.98</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.89</td>
</tr>
<tr>
<td>Houston</td>
<td>0.98</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td>St. Louis</td>
<td>0.98</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>Phoenix</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.04</td>
</tr>
<tr>
<td>Dallas</td>
<td>0.98</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.06</td>
</tr>
<tr>
<td>New York</td>
<td>1.00</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>1.01</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>1.03</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>1.03</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>1.03</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>1.03</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>1.05</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>1.05</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>1.06</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>1.13</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>1.15</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>1.22</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>0.83</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>0.86</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>0.87</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.93</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>1.06</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>1.01</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>1.05</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>1.04</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>1.07</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>1.09</td>
<td>1.19</td>
</tr>
<tr>
<td></td>
<td>1.16</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>1.10</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>0.83</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>-0.11</td>
</tr>
</tbody>
</table>
Table A.6: Cities Ranked According to Grocery Costs Calculated Using Homothetic and Non-Homothetic Models

<table>
<thead>
<tr>
<th>Market</th>
<th>Homothetic</th>
<th>Non-Homothetic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$16,896</td>
<td>$93,411</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Washington, DC-Baltimore</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Atlanta</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Detroit</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>Columbus</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>Tampa</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>Houston</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>St. Louis</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>Phoenix</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Dallas</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>New York</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Miami</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Boston</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Charlotte</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Denver</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Buffalo-Rochester</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>Seattle</td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td>Chicago</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>San Antonio</td>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td>Sacramento</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>San Francisco</td>
<td>23</td>
<td>4</td>
</tr>
</tbody>
</table>
Table A.7: Cross-City Variation in Household Demographics within Income Groups

Panel A: OLS Regression

<table>
<thead>
<tr>
<th></th>
<th>Ln(Household Size(_{h,c}))</th>
<th>Ln(Age(_{h,c}))</th>
<th>Ln(Education(_{h,c}))</th>
<th>Ln(Per Capita Income(_{c}))</th>
<th>Ln(Household Income(_{h}))*</th>
<th>Constant</th>
<th>Observations</th>
<th>R-Squared</th>
<th>Household Income Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Per Capita Income(_{c}))</td>
<td>-0.003</td>
<td>0.025</td>
<td>-0.026</td>
<td>0.233</td>
<td>0.012</td>
<td>0.144</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.004]</td>
<td>[0.083]</td>
<td>[0.032]</td>
<td>[0.199]</td>
<td>[0.014]</td>
<td>[0.133]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(Household Income(_{h}))*</td>
<td>-</td>
<td>-0.003</td>
<td>-</td>
<td>-0.024</td>
<td>-</td>
<td>-0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.008]</td>
<td></td>
<td>[0.018]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.132***</td>
<td>1.089***</td>
<td>4.182***</td>
<td>3.784***</td>
<td>2.418***</td>
<td>2.217***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.046]</td>
<td>[0.124]</td>
<td>[0.333]</td>
<td>[0.451]</td>
<td>[0.145]</td>
<td>[0.302]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.968</td>
<td>0.968</td>
<td>0.187</td>
<td>0.187</td>
<td>0.202</td>
<td>0.202</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Income Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Logit Regression

<table>
<thead>
<tr>
<th></th>
<th>I(Married(_{h,c}))</th>
<th>I(White(_{h,c}))</th>
<th>I(Hispanic(_{h,c}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Per Capita Income(_{c}))</td>
<td>-1.175***</td>
<td>0.227</td>
<td>-0.996</td>
</tr>
<tr>
<td></td>
<td>[0.265]</td>
<td>[1.964]</td>
<td>[0.799]</td>
</tr>
<tr>
<td>Ln(Household Income(_{h}))*</td>
<td>-</td>
<td>-0.133</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[0.172]</td>
<td></td>
<td>[0.414]</td>
</tr>
<tr>
<td>Constant</td>
<td>12.141***</td>
<td>10.091***</td>
<td>11.896</td>
</tr>
<tr>
<td></td>
<td>[2.762]</td>
<td>[4.690]</td>
<td>[8.387]</td>
</tr>
<tr>
<td>(Pseudo) R-squared</td>
<td>0.458</td>
<td>0.458</td>
<td>0.0229</td>
</tr>
<tr>
<td>Household Income Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1; Robust standard errors, clustered by city, in brackets.

Notes:
1. Age and education refer to the mean age and education of one or two heads of household, respectively; both are measured in years.
2. This sample includes all Nielsen households sampled from 2003 to 2005 with non-missing demographic information reporting an aggregate annual income above $10,000.
3. Household income is adjusted for household size, as described in Section 2.