Essays on Prices and Product Variety Across Cities

Jessie Handbury

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

Abstract

Essays on Prices and Variety Across Cities

Jessie Handbury

Urban economists are fundamentally concerned with the distribution of economic activity across cities and the resulting variation in urban welfare. The central equilibrium concept is that mobile individuals choose their locations optimally. If equivalent individuals choose two different locations, they must be spatially indifferent. For a system of cities to be at a spatial equilibrium, real income must be equalized across space. The purchasing power residents enjoy in different cities is, therefore, central to urban economics. It is key to determining both the distribution of economic activity, through the indifference condition, and consumer welfare at this equilibrium.

The New Economic Geography (NEG) literature initiated by Krugman (1991) highlights the role of pecuniary externalities in urban economics. Net of housing and other congestion costs, consumers prefer larger cities because they have lower average prices and more varieties. These price and variety differences impact real incomes directly, by lowering the consumer price index or indirectly, by lowering the producer price index, improving firm productivity, and nominal wages. My dissertation focuses on the direct consumer channel. Specifically, I study the differences in consumer prices and variety across U.S. cities, with a view towards measuring the consumption externalities that drive the agglomeration that determines the relative sizes and demographics of these urban centers.

In Krugman (1991), agglomeration is driven by consumption externalities related to differences in the average price across locations. Variety plays a central role here: the price index is lower in larger cities because more varieties are produced there, but all products are available everywhere, so there are no direct consumption gains from variety. Chapter 1 extends this model to allow for these variety differences across cities via an extensive margin of intercity trade. In this model, scale economies yield more production variety in larger cities so that, with positive trade costs, consumers in these locations can benefit from two consumption advantages - lower average prices and more varieties of products. This model sets the stage for the empirical chapters that measure the impact of these consumption benefits on consumer utility.

The urban economics literature has paid relatively little attention to these consumption externalities for two reasons. First, the fact that wages are higher in larger cities indicates, in the context of a spatial equilibrium model, that purchasing power must be lower in these locations, which is inconsistent with the NEG theoretical predictions that price indexes are lower in larger cities. Second, there are reasons to believe that intranational trade frictions are much smaller than those across international borders. While there is a large body of work documenting that these frictions yield significant variation in the prices charged and varieties offered across different countries, we know relatively little about the spatial variation of price and variety in a domestic context. Chapter 2, co-authored with David E. Weinstein, addresses both of these issues. We first show that, controlling for purchaser demographics and store amenities, the prices of tradable products do, in fact, vary across cities as predicted by Krugman (1991). Additionally, we find that there are significant differences in the variety of products offered in large as opposed to small cities. We finally measure the extent to which this variation in prices and variety lowers price indexes for tradable products in large cities relative to small cities. This low price index over tradables in large cities is consistent with nominal wages being higher there, as long as there are congestion costs that equalize real income across locations.

A major assumption in the work described above is that the consumption benefits of cities do not vary systematically across consumer types. The final chapter of my dissertation considers how systematic differences in prices and variety across space might impact different individuals differently. Previous research has tested whether firms vary prices and product offerings in order to cater local tastes in a market. I extend this analysis to structurally estimate how this behavior of individual firms differentially affects the price indexes faced by consumers with systematically heterogeneous tastes. I allow for tastes to vary with income and find that poor consumers face lower costs in low-income cities, while the opposite holds for wealthy consumers, whose tastes are better-suited to the variety of products available in high-income cities.

In conclusion, this dissertation finds that there is significant variation in prices and variety across U.S. cities. In particular, the spatial distributions of prices, variety, and consumers in the U.S. are correlated in a manner consistent with there being consumption externalities. Consumers in larger cities have access to more varieties of products at a

lower average price and, conditional on city size, consumers have access to varieties better-suited to the tastes of they share with their income group in cities with per capita incomes closer to their own. These patterns are reflected in the variation in the tradables price indexes that consumers face across large and small and across wealthy and poor cities. This evidence on the existence of the pecuniary consumption externalities hypothesized in early NEG papers supports a growing literature exploring the role of consmption externalities in generating the aggregate and skill-biased agglomeration patterns that the more recent NEG literature associates with production externalities.

Contents

List of Figures

iii

iv

List of Tables

Acknowledgments

I would like to sincerely thank my advisors David Weinstein and Kate Ho.

I also want to thank the many faculty and colleagues for advice and guidance over the course of my graduate career: Chris Conlon, Don Davis, Jonathan Dingel, Ron Findlay, Amit Khandelwal, Joan Monras, Emi Nakamura, Mike Riordan, Bernard Salanié, Catherine Thomas, Eric Verhoogen, and Jonathan Vogel.

For supporting my research, I thank Columbia University's Program for Economic Research. I also thank the United States Department of Agriculture for providing access to the Nielsen HomeScan data used in this research.

Dissertation Committee

Chris Conlon
Columbia University
Department of Economics

Katherine Ho
Columbia University
Department of Economics

Amit Khandelwal
Columbia University
Graduate School of Business

Jonathan E. Vogel
Columbia University
Department of Economics

David E. Weinstein
Columbia University
Department of Economics

Chapter 1

Prices and Variety in a New Economic Geography Model

Jessie Handbury

1.1 Introduction

Consumption externalities are the central agglomerating force in Krugman (1991) and continue to play a key role in the New Economic Geography (NEG) literature that has followed. These externalities arise from a very simple structure based on trade costs and scale economies. Trade costs are positive between cities, so locally-produced goods must be cheaper in the local market than they are elsewhere. Fixed costs of production generate scale economies under which larger cities will produce more varieties than smaller ones. Consumers in large cities, therefore, purchase more locally-produced goods and, therefore, face a lower average price.

This paper studies a complementary and related source of consumption externalities: differences in variety across space. It has been widely documented that not all varieties of products are traded between countries,[1] limiting the access consumers have to varieties that are not produced locally. If domestic trade flows are also characterized by this extensive margin variation, cross-city differences in the variety produced will be reflected in cross-city differences in the variety available for consumption. When not all products

---

[1]See, for example, Hummels and Klenow (2005) and Bernard, Jensen, Redding, and Schott (2007).

are traded, consumers in large cities will have access to more varieties of goods than consumers in small cities. These consumption variety differences will tend to lower the price index in large cities relative to small cities, yielding an agglomeration mechanism that is complementary to the average price mechanism described above. The tradables price index will be lower in large cities because consumers there have access to more locally-produced products and more products in general.

I present a two-city model that yields asymmetric spatial equilibria in which some, but not all, of the firms in each location export their products. I show that, when some but not all products are traded, price and variety mechanisms work simultanously to yield a lower price level in the larger city. This model extends the traditional Krugman (1991) model in a number of dimensions that have been explored to some extent in other papers. First, I require consumers to purchase housing in the city where they reside. The starting point for this assumption is Helpman (1998), a variant of the Krugman (1991) model in which housing costs, rather than an agricultural sector, act as the dispersion force causing the real wages to equilibrate across locations. Helpman's model is similar to Krugman (1991), however, in that it can only produce the equilibrium result that larger cities have more varieties under infinite trade costs. To overcome this limitation, I also allow for firm-specific fixed costs of exporting between cities. When only a subset of firms can earn profits from exporting, the number of varieties varies across locations, much in the spirit of Melitz (2003). As in the original Krugman (1991) model, more locally-produced varieties are available in the larger city, so the average price is lower there also. These variety and price effects combine to yield a lower non-housing price index in the large city relative to the small city.

Models featuring quadratic linear demand generate similar predictions on the variation of price and variety across markets. In Ottaviano, Tabuchi, and Thisse (2002), for example, more varieties are produced in larger markets and, under positive trade costs, this results in more intense competition and lower mark-ups in these locations. Prices are, on average, lower in larger cities both because consumers pay (a fraction of) transport costs on fewer varieties and because mark-ups are lower on all varieties. Behrens and Robert-Nicoud (2011) introduce firm heterogeneity to a quadratic linear demand-based model generating the additional prediction that more varieties are both produced and consumed

in larger cities. These authors show that larger cities will have both lower prices and more varieties of traded goods products available, but the role of price and variety as an agglomeration force is limited in their framework because consumers are not free to move between differentiated production centers.[2]

By contrast, price and variety are both central to agglomeration in the model presented below. In this heterogeneous-firms extension of Helpman's CES-based geography model, consumers are attracted to the larger city because of the lower traded goods prices and greater variety availability there. However, as the population of one city grows, its rental rate also rises. I show below that this can result in an asymmetric spatial equilibrium in which both cities are populated and the large city has a lower price index for tradable goods but higher rental prices. At these long-run equilibria, the lower cost of tradable goods is exactly offset by higher rents, such that real wages are equalized across the two locations and consumers have no incentive to move.

The simple model presented below is designed to highlight the agglomerating role of consumption externalities. There are two cities with identical and fixed stocks of housing. In the short-run, identical consumers reside in one of the two cities, working in the production of a freely-traded intermediate good and earning rents from their share of the economy-wide stock of housing. These assumptions imply that a consumer's nominal income is the same regardless of where they live; their utility, or standard of living, in a city depends only the purchasing power they have there. Consumers spend their income on housing and the differentiated products available where they live, including both locally-produced and, if available, imported varieties. People in more populated regions face higher housing costs, but lower price indexes over differentiated products. The high housing costs in large cities is due to inelastic housing supply and drives down the local standard of living in these locations. The differentiated price index is lower in large cities because there are more varieties available in these cities at a lower average price. The cost of housing and the differentiated product price index work in opposite directions in determining the relative standard of living across the two locations.

I solve the model to generate predictions for the relative cost of housing and differ-

---

[2]In Behrens and Robert-Nicoud (2011), people can move between the agricultural and urban economies within each region, but not between regions.

entiated product price index across two cities as a function of their relative size. The relative cost of housing can be solved for analytically and is proportionate to city size. The differentiated product price indexes in the two locations only vary due to the number and export decisions of final goods producers.[3] Specifically, the differentiated product price index is decreasing in both the number of varieties available in a location and the share of these varieties that are produced locally. I will refer to the first of these effects as the "variety effect" and the second as the "price effect," since the larger the share of varieties consumed locally that are also produced locally, the lower the average price of differentiated products in the city. The variety effect is generated on the extensive margin of variety in that it works to lower the relative price index faced in cities that have a greater total number of varieties available. The price effect is generated on the intensive margin in that it works to lower the relative price index faced in cities where a greater share of those varieties available are produced locally.

The existence of the price and variety effects depend on the number and the export decisions of final goods producers. The number of final goods producers in each location, all of which sell on the local market, is increasing in city size when there are positive marginal trade costs. These costs take the "iceberg" form, such that a certain proportion of the quantity exported by the producer arrives in the export market for sale to the consumer. These costs lower the potential profits firms can earn per consumer in the export market relative to the domestic market. Holding the number of firms equal across the two cities, total firm profits will be higher in the larger city, generating more firm entry into (or less firm exit from) the larger city until profits are equal zero in both locations. More populated cities will therefore be able to sustain production of a larger number of differentiated products. The difference in the mass of firms locating in the large relative to the small city is the main driver behind the price effect. It turns out that, at the interior equilibria, the share of firms that export from each city is identical across the two cities. This implies that the share of consumption varieties that are produced locally is increasing in the total number of varieties produced locally. Thus, the economics of scale

---

[3]The differentiated product price index is determined by the prices charged by, and the number and export decisions of, final goods producers in each city. These producers face identical marginal production and trade costs and demand curves across locations and, therefore, charge an identical price for all domestic varieties and another, higher, identical price for all imported varieties of differentiated products.

that yield more varieties are produced in the larger city, also yield the intensive margin price effects that lower the average differentiated product price, and also tend to lower the differentiated price index, in the larger city when there are transpost costs on these products.

Where these transport costs impact the prices that firms set in export markets, their decisions over whether to export or not are governed by a firm-specific fixed cost of export. Firms will export if their ex-post export profits are sufficient to cover this cost, drawn from a random distribution.[4] Firms face greater demand for differentiated products in the larger cities, so will be more likely to cover their fixed export cost with export sales to a large city than to a small city. Larger cities will always import all of the varieties that smaller cities do and, as such, more products will be available in larger cities than in smaller cities. This difference in total product availability generates the variety effect.

Since larger cities have a greater total number of varieties are available as well as a greater share of the available varieties is locally-produced implies that both the price and variety effects work to lower the relative price index there when some, but not all, firms are exporting from both locations. Although these "interior" equilibria are the focus of this paper, the model presented here also nests the Krugman (1991) model in the sense that there exist two types of "corner" equilibria in which all and no firms export, respectively. If, for example, the profits that firms earn from exporting are lower than the lower bound of the cost distribution then no firms will export. In this type of equilibrium, as in Krugman (1991), there will be no "price" effect: only local varieties will be available in each location, making the share of local varieties in consumption equal to one in both locations. If, alternatively the profits that firms earn from exporting are greater than the upper bound of the cost distribution then all firms will export. In this case, there will be no "variety" effect: all varieties will be available in both locations. If, however, export

---

[4]An alternative way to achieve selection into exporting is with a fixed export cost and heterogeneity in firm productivity and, therefore, ex-post export profits, as in Melitz (2003). I find that heterogeneous fixed export costs is the simplest way to generate an extensive margin of trade in keeping with the original Krugman (1991) model in that variation in the differentiated-product price indexes between two locations is driven by asymmetries in the number of firms located in each and the share of these firms that export. Heterogeneous firm productivities would instead introduce additional sources of variation in the consumer price index. For example, imported product prices will tend to be lower in the smaller city because the productivity cutoff for big-city firms to export to the small city will be lower than the productivity cutoff for small-city firms to export to the large city. Exploring these additional asymmetries is beyond the scope of this paper.

profits are within the cost distribution, then an intermediate number of firms will make their varieties available in both cities and both the price and variety effects will impact how the differentiated price index varies with city size.

In the next section, I present the key features of the model. Section 3 defines the equilibrium conditions for the economy, and shows that both in the short-run, where prices and variety are determined taking the allocation of people between cities as given, and in the long-run, where consumers are free to move between locations. Section 4 presents the main results of the model showing that the price index over differentiated products is lower in the larger city due to both variety and price effects for any short-run equilibrium in which these products are produced and exported from both cities. Section 5 concludes.

1.2   Basic Model

I begin by assuming that there are L total consumers allocated across a set of two cities denoted $\mathbb{I} = \{1, 2\}$, such that L_i people live in each city i, and $L_1 + L_2 = L$. There are three sectors: housing, intermediate good production, and final good production. A fixed and equal endowment of land in each location is rented to consumers for housing. Revenues from land rentals are shared equally across all consumers. Consumers also get income from working in the intermediate goods sector. Intermediate goods are freely-traded and sold to final goods producers in a competitive market. Firms in the final goods sector purchase technologies that enable them to convert intermediate goods into differentiated varieties. Producers can trade their final goods by paying a firm-specific fixed export cost as well as iceberg transport costs.

The starting point for this model is Helpman (1998). Before going into the details of the model, I will highlight three major departures I take from this model. The first are the heterogeneous fixed costs of exporting from one city to another. These costs generate the central result that more varieties of goods are available in the larger city. The second is that differentiated goods producers use a freely-traded intermediate input in place of labor. This implies that wages and the mill price of differentiated products are equal across the two cities, even when the prices of differentiated products are not. This assumption is

very helpful in producing analytic, as opposed to numeric solutions to the model.[5] Finally, I add a minimum housing requirement which serves as an anti-agglomeration force that prevents the entire population from ending up in one city when the symmetric equilibrium is unstable.

1.2.1 Consumption

I assume that consumers in each city consume non-tradable housing services and tradable final goods. Preferences are Cobb-Douglas over any housing, H, consumed above and beyond a minimum housing requirement, β, and an aggregate final good bundle, X. The aggregate final good, X, represents the constant elasticity of substitution (CES) utility that consumers attain from consumption of the final good varieties. Each final good variety is denoted by u, and residents of city i can purchase any of the set of varieties sold in city i, denoted $\mathbb{U}_i$. A city i resident facing rents r_i and final goods prices, $p_i(u)$, allocates their income I over housing and final good varieties to maximize utility as follows:

$$\max_{\left\{H>0,\{x(u)\}_{u\in\mathbb{U}_i}\right\}} X^\alpha (H-\beta)^{1-\alpha} \text{ subject to } \int_{u\in\mathbb{U}_i} p_i(u)x(u)du + r_iH \leq I \text{ where } X = \left[\int_{u\in\mathbb{U}_i} x(u)^\rho du\right]^{1/\rho}$$

where $\sigma = 1/(1-\rho)$ is the elasticity of substitution across final good varieties. The solution to the above problem implies the following individual consumer demand curves for housing and final goods in city i:

$$H_i = \frac{(1-\alpha)(I-\beta r_i)}{r_i} + \beta \tag{1.1}$$

$$X_i = \frac{\alpha(I-\beta r_i)}{P_i} \tag{1.2}$$

$$x_i(u) = X_i \left(\frac{p_i(u)}{P_i}\right)^{-\sigma} \quad \forall u \in \mathbb{U}_i \tag{1.3}$$

---

[5]Both Krugman (1991) and Helpman (1998) allow for wages to vary with city size. When relative wages increase in relative city size, equilibrium housing rents increase non-linearly with city size providing the extra anti-agglomeration force necessary to yield asymmetric equilibria in which both cities are populated. When fixed export costs are added to the original Helpman (1998) model without any further adjustments, we are able to show the existence of stable asymmetric equilibria in which both cities are populated and there are more varieties available at a lower average price in the larger city, but the variation in relative wages makes it impossible to prove this result generally.

$$\text{where } P_i = \left[\int_{u \in \mathbb{U}_i} p_i(u)^{1-\sigma} du \right]^{\frac{1}{1-\sigma}}$$

Consumers are perfectly mobile across cities, so their locational choices maximize their location-specific indirect utility:

$$V(I, P_i, r_i) = \frac{k(I - \beta r_i)}{P_i^{\alpha} r_i^{1-\alpha}} \tag{1.4}$$

where k is equal to a constant function of parameter values and $(I - \beta r_i)$ is the income in excess of that needed to meet the minimum housing requirement.[6] Holding income fixed, equation (1.4) tells us how the price index for differentiated goods, P_i, and housing rents, r_i, affect utility in each city. The next task is to put some more structure on the model so that we can understand how these price indices move with city size.

1.2.2   Income-generating Activities

I assume that each consumer is endowed with one unit of labor and an even share of the land in each location. The income, I, generated from these endowments is equal to the sum of the wage, w, and rental income, denoted $\tilde{r}$:

$$I = w + \tilde{r} \tag{1.5}$$

Wages are earned in the production of tradable intermediate goods. Labor is the single input in the production of the homogeneous intermediate good. The intermediate good can be produced in every location using the same constant returns technology, with a unit labor requirement normalized to equal one. It is freely traded across locations and sold to final goods producers on a competitive market at a price, p^0, which we also normalize to one. This implies that the competitive wage rate equals one in all locations:

$$w = 1 \tag{1.6}$$

---

[6]It is possible to show that, in equilibrium, $(I - \beta r_i) > 0$.

Rental income is earned in the housing market. Each city has a common fixed allocation of land, $\bar{H}$, rented by consumers for housing at a city-specific rental rate, r_i.

The rental rate r_i in city i is determined in equilibrium to clear the city's land market, equalizing the supply of the available land, $\bar{H}$, with the demand of the city's L_i residents:

$$\bar{H} = H_i L_i$$

Using the demand curve for H_i provided in equation (1.1), the land-clearing condition above implies that the market-clearing rental rate is

$$r_i = (1 - \alpha) I L_i \Phi_i^{-1} \tag{1.7}$$

where $\Phi_i = \bar{H} - \alpha\beta L_i$.

The total rental income in city i is equal to $r_i \bar{H}$. The sum of this rental income across all cities is evenly distributed across consumers who each receive

$$\tilde{r} = \frac{\bar{H}(r_1 + r_2)}{L} = \frac{\bar{H}I(1 - \alpha)(L_1/\Phi_1 + L_2/\Phi_2)}{L} \tag{1.8}$$

I can now use (1.6), and (1.8) to solve for consumer income:

$$I = w + \tilde{r} = \left[1 - \frac{\bar{H}(1 - \alpha)(L_1/\Phi_1 + L_2/\Phi_2)}{L}\right]^{-1} \tag{1.9}$$

Equations (1.7) and (1.9) together characterize a fixed point for consumer income and for rents in each city $i \in \mathbb{I}$.

1.2.3 Final Good Production

The final goods market is monopolistically competitive. Any consumer in a location can decide to start a firm by paying a fixed cost of f units of the intermediate good.[7] This fixed cost can be thought of as the intermediate input required to develop a technology

---

[7] There is free entry into final goods production so, in equilibrium, expected profits from entry into production in a given city are zero. Some consumers will make positive profits and others will make losses such that entrepreneurial activity is not a source of consumer income at the city level. I allow consumers to start more than one firm so that the number of firms that can be started in a location is not limited by the number of consumers residing there. However, I assume that firms owned by the same consumer do not collude in their pricing strategy and act as two separate firms making independent pricing decisions.

that the entrepreneur can in turn use in final goods production. This technology costlessly transforms further units of the intermediate good into differentiated final good on a one-to-one basis. Once the firm has developed this technology, its marginal cost of production is therefore equal to the price of the intermediate input, $p^0 = 1$.

Equations (1.2) and (1.3) together define the quantity, $x_i(u)$, that each of the L_i consumers facing an aggregate price level of P_i in a city i will demand of a final good variety u sold at a price $p(u)$. I assume that differentiated product firms ignore the impact of their pricing behavior on the aggregate price index. Firms, therefore, face a constant price elasticity, $1 - \sigma$. Since each firm is a monopolist in the production of its variety, u, firms follow an inverse elasticity pricing rule, charging a fixed mark-up, $\sigma/(\sigma - 1) = 1/\rho$, on their marginal costs.

All firms that produce final goods sell these goods on their local retail market. A city i firm's marginal costs for domestic sales consist of the cost of one unit of the intermediate good, $p^0 = 1$:

$$c_{ii} = 1$$

Firms can also export their products to other cities for a product-specific fixed export cost, $f_x(u)$. Firms export decisions will be discussed in more detail below. Firms that decide to export their product pay "iceberg" transport costs in that they must ship τ units of their final good in order to sell a single unit in the exporting market. The marginal cost of producing and shipping one unit of a product in city j is therefore equal to the cost of τ units of the intermediate input, $\tau p^0 = \tau$:

$$c_{ij} = \tau$$

Denote the price that city i firms charge in city j as p_{ij}. This optimal price can be summarized as

$$p_{ij} = \frac{c_{ij}}{\rho} \text{ where } c_{ij} = \begin{cases} 1 & \text{if } i = j \\ \tau & \text{otherwise} \end{cases} \tag{1.10}$$

Since all domestically-produced final goods are sold at an identical price and all imported final goods are sold at the same location-specific price, we can write the final goods

price index in city j as

$$P_j = \left(\sum_{i \in \mathbb{I}} n_{ij} p_{ij}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} = \frac{1}{\rho} \left(\sum_{i \in \mathbb{I}} n_{ij} c_{ij}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{1.11}$$

where n_{ij} is the number of varieties that are imported to city i from city j and n_{jj} is the number of firms producing locally in city j. I can now use (1.2) and (1.3) to calculate the sales of a city i firm makes to each consumer in city j:

$$x_{ij} = \frac{\alpha(I - \beta r_i)\rho}{c_{ij}^{\sigma} \left(\sum_{k \in \mathbb{I}} n_{kj} c_{kj}^{1-\sigma} \right)} \tag{1.12}$$

The ex-post profits for a firm in city i from sales in city j are, therefore,

$$\pi_{ij} = \frac{L_j x_{ij} c_{ij}}{\sigma - 1} \tag{1.13}$$

1.2.4 Zero-Cutoff Profit Condition

While all firms will produce for their local market, only firms whose ex-post profits are greater than their fixed export cost draw, $f_x(u)$, will export to a given city j. Assuming that the fixed costs of export are drawn with replacement from a uniform distribution on $[\underline{f_X}, \overline{f_X}]$, the share of city i firms whose fixed export costs are low enough to yield ex-ante profits from exporting to location j will be

$$\frac{n_{ij}}{n_{ii}} = \Pr\left[\pi_{ij} \geq f_x(u) \right] = \Pr\left[f_x(u) \leq f_{ij}^* \right] = \frac{f_{ij}^* - \underline{f_X}}{\overline{f_X} - \underline{f_X}} \tag{1.14}$$

where $f_{ij}^* = \max\left(\underline{f_X}, \min\left(\overline{f_X}, \pi_{ij} \right) \right)$ denotes the amount of fixed export costs that could be potentially offset by ex-post export profits.

1.2.5 Free Entry Condition

Free entry in production requires that, in equilibrium, expected profits are zero. Expected profits are equal to the sum of local profits and the expected value of positive export profits, less the fixed cost of entry, f. The free entry condition for firms in city i is,

therefore,

$$\pi_{ii} + \Pr\left[f_x(u) \leq \pi_{ij}\right]\left(\pi_{ij} - \mathbb{E}\left[f_x(u)|f_x(u) \leq \pi_{ij}\right]\right) = f$$

where $j \neq i$. The second term in this equation is the expected profits from exporting. By Bayes's rule, the expected export profits are equal to the probability of exporting, $\Pr\left[f_x(u) \leq \pi_{ij}\right]$, multiplied by the expected profits from exporting conditional on exporting. Since the *ex-post* profits from exporting are identical across firms, the export profits vary only with the random fixed costs of export and the conditional expected export profits are equal to the ex-post profits from exporting, π_{ij}, less the expected fixed cost of export conditional on exporting, $\mathbb{E}\left[f_x(u)|f_x(u) \leq \pi_{ij}\right]$. The unconditional expected export profits term can be simplified using the zero-cutoff profit condition as follows:

$$
\begin{aligned}
\Pr\left[f_x(u) \leq \pi_{ij}\right]\left(\pi_{ij} - \mathbb{E}\left[f_x(u)|f_x(u) \leq \pi_{ij}\right]\right) &= \Pr\left[f_x(u) \leq f_{ij}^*\right]\left(\pi_{ij} - \mathbb{E}\left[f_x(u)|f_x(u) \leq f_{ij}^*\right]\right) \\
&= \left(\frac{f_{ij}^* - \underline{f_X}}{\overline{f_X} - \underline{f_X}}\right)\left(\pi_{ij} - \frac{\underline{f_X} + f_{ij}^*}{2}\right)
\end{aligned}
$$

The free entry condition in city i is, therefore,

$$f = \pi_{ii} + \left(\frac{f_{ij}^* - \underline{f_X}}{\overline{f_X} - \underline{f_X}}\right)\left(\pi_{ij} - \frac{\underline{f_X} + f_{ij}^*}{2}\right) \tag{1.15}$$

where $j \neq i$.

1.3   Equilibrium

This section develops the conditions for equilibrium in our spatial economy. In the long run, consumers optimally choose their city. In the short run, consumers choose their consumption of housing and final goods varieties. Consumers can also act as entrepreneurs. In this role, they choose whether to start one or more firms in a city. All firms sell their product locally, but each chooses whether to export their variety to other cities. Firms also set prices for their variety in each market where it is sold.

1.3.1 Short-Run Equilibrium

I first define a short-run equilibrium for a system of cities in which a population L is distributed across two cities, $\mathbb{I} = \{1, 2\}$, which with a land endowment $\bar{H}$, such that $\sum_{i \in \mathbb{I}} L_i = L$. This short run equilibrium is defined as a set of differentiated product prices $\mathbb{P} = \{p_{ij}\}_{i,j \in \mathbb{I}}$, quantities $\mathbb{X} = \{x_{ij}\}_{i,j \in \mathbb{I}}$, and product counts $\mathbb{N} = \{n_{ij}\}_{i,j \in \mathbb{I}}$, as well as, housing rents $\mathbb{R} = \{r_i\}_{i \in \mathbb{I}}$ and quantities $\mathbb{H} = \{H_i\}_{i \in \mathbb{I}}$ such that consumers and firms optimize and markets clear.

The conditions under which the housing market clears were derived in Section 1.2.2. Equations (1.7) and (1.9) together characterize a fixed point for consumer income and for rents in each city $i \in \mathbb{I}$ that clear the housing market in each city. At this income and these rents, residents in each city optimally choose their consumption of housing according to (1.1).

The conditions for equilibrium in the final goods market were derived in Section 1.2.3. Equation (1.10) summarizes the profit-maximizing prices set by differentiated product firms. The quantities that equalize demand and supply of differentiated products in each location are defined, conditional on income, rents, and the number of locally-produced and imported products sold in each market, in equation (1.12). There is free entry into final good production, so entrepreneurs will optimally purchase the technology required to start producing a differentiated product as long as profits are positive. The free entry condition, defined in equation (1.15) for each city i, implicitly defines the number of firms that exist when expected profits from entry are zero. Similarly, the zero profit cutoff, defined in equation (1.14), implicitly defines the share of firms that export from each city i to the other city $j \neq i$ such that the marginal exporting firm earns zero profit from export sales after paying their fixed cost of export.

The intermediate goods and labor markets clear by Walras's Law and at equilibrium, inter-city markets are balanced by international transfers of housing rents.

1.3.2 Long-Run Equilibrium

The previous section characterized the short-run equilibrium for the system of cities, in which I assume that consumers' do not make locational choices and, therefore, the pop-

ulation of each city, $\mathbb{L} = \{L_i\}_{i \in \mathbb{I}}$, is fixed. In the long-run, consumers are mobile and optimally choose their location in order to maximize their indirect utility from housing and final goods consumption, $V(I, P_i, r_i)$, defined in equation (1.4). The system of cities will be in long-run equilibrium when the short-run equilibrium conditions hold and either: utility is equalized across locations, such that no consumer can make utility gains from moving between cities, or utility is higher in one location and all consumers live in this location.

1.4   Existence of Long-Run Equilibrium

There are three classes of long-run equilibria: symmetric equilibria with identical city sizes, asymmetric equilibria in which one location is not populated, and asymmetric equilibria in which both are populated to differing degrees. The goal of this theory is to bring Krugman (1991)'s model closer to reality for the purpose of empirical testing. Since this empirical work will be based on data from non-empty cities with varying populations, we will focus on the predictions of the model for the third class of long-run equilibria. In particular, our focus will be on the subset of long-run asymmetric equilibria for which there is differentiated production in both cities and some, but not all, products are traded in both directions.

The existence of these equilibria is stated in the following proposition:

Proposition 1. *For certain parameter values, stable, asymmetric equilibria exist in which the two cities have positive populations; differentiated products are produced in both cities; and some, but not all, of these products are exported from each city.*

Proof. The system of cities will be in long-run equilibrium when utility is equalized across the two locations, such that no consumer can make utility gains from moving between cities, or when utility is higher in one location and all consumers live in this location. The dynamics of the system of cities, and its long-run equilibria, will be characterized by how the relative utility across two locations varies with allocation of consumers across those populations. Without loss of generality, we will consider the utility of consumers in city i relative to city j, where $i \in \mathbb{I} = \{1, 2\}$ and j denotes the other city in the pair. We will measure the allocation of consumers across the two locations using the share

of the population in city i, $\lambda = L_i/L$. The short-run equilibrium conditions in Section 1.3.1 and equation (1.4) together define this relative utility as a function of population allocation, λ, and model parameters, Θ. Below we will characterize how the equilibrium relative utility varies across short-run equilibria with different population allocations, λ, differently for different sets of parameter values Θ. This analysis reveals a set of parameter values under which the empirically-relevant asymmetric population allocation is a stable, long-run equilibrium for the system.

First, let $V_i(\lambda, \Theta)$ denote the indirect utility of a consumer in city i at the short-run equilibrium defined for population allocation λ and parameters Θ. Equation 1.4 implies that this indirect utility can be written as

$$V_i(\lambda, \Theta) \;\; = \;\; \frac{k(I(\lambda, \Theta) - \beta r_i(\lambda, \Theta))}{P_i(\lambda, \Theta)^\alpha r_i(\lambda, \Theta)^{1-\alpha}}$$

where $I(\lambda, \Theta)$, $r_i(\lambda, \Theta)$, and $P_i(\lambda, \Theta)$ also each take their short-run equilibrium values, conditional on population allocation, λ, and parameters, Θ. We can now write the log of the utility in city i relative to city j as

$$\hat{V}(\lambda, \Theta) \;\; = \;\; \hat{D}(\lambda, \Theta) - \alpha \hat{P}(\lambda, \Theta) - (1-\alpha)\hat{r}(\lambda, \Theta)$$

where $D_i(\lambda, \Theta) = I(\lambda, \Theta) - \beta r_i(\lambda, \Theta)$ denotes a city i consumer's disposable income after paying for their minimum housing requirement at the (λ, Θ)-short run equilibrium. Positive values of $\hat{V}(\lambda, \Theta)$ imply that consumers utility is greater in city i than in city j yielding agglomeration when city i is already larger than city j, or $\lambda > 0.5$, and dispersion when city i is smaller than city j, or $\lambda < 0.5$.

An equilibrium will exist at a population allocation λ if $\hat{V}(\lambda, \Theta) = 0$, or $\alpha \hat{P}(\lambda, \Theta) = \hat{D}(\lambda, \Theta) - (1-\alpha)\hat{r}(\lambda, \Theta)$. The population allocations for which this is the case depends crucially on parameters Θ. Our object here is to show the existence of parameter values, Θ, under which the system yields long-run equilibrium for some $\lambda \in (0,1)$ that is not equal to 0.5, *i.e.* a long-run equilibrium in which both cities are populated, but not equally. We first show that such an asymmetric long-run equilibrium exists for one set of parameter values.

Figure 1.1 presents the relative utility in city i relative to city j as the share of consumers living in city i, λ, increases from 0.25 to 0.75, when $\alpha = 0.9$, $\sigma = 2$, $\tau = 1.5$, $\underline{f_X} = 0$, $\overline{f_X} = 1$, and $\beta = 0.01175$. This chart shows that, for these parameter values, the relative utility is equalized across locations at three population allocations: $\lambda = 0.5$, $\lambda \approx 0.33$ and $\lambda \approx 0.66$. There exists, therefore, three equilibria for the system of cities at these parameter values, one symmetric and two asymmetric. While our primary focus is on showing the existence of the asymmetric long-run equilibria, it is interesting to note that these equilibria are both stable, while the symmetric equilibrium is unstable. Suppose that an exogenous shock moved $\Delta\lambda$ of the total population from city j to city i, increasing the population allocation from 0.5 to $0.5 + 2\Delta\lambda$. Since relative utility is increasing in λ at the symmetric equilibrium, *i.e.* $\partial\hat{V}(0.5, \Theta)/\partial\lambda > 0$, the new population allocation will yield higher indirect utility to consumers in city i than in city j, encouraging further migration from city j to city i. Similarly, an exogenous shock that moved $\Delta\lambda$ of the total population from city i to city j would increase the utility in city j relative to city i encouraging further consumers follow. Conversely, the asymmetric equilibria are stable because relative utility is decreasing in λ at these points. If λ was at one of these points and an exogenous shock were to move consumers from city j to city i, the indirect utility of consumers living city j would decrease relative to city i, encouraging consumers to move from city i to city j whereby reversing the original migration.

We now show that, at the asymmetric equilibria described above, differentiated products are produced in both cities; and some, but not all, of these products are traded in both directions. The upper left panel in Figure 1.2 shows that there are approximately 30 varieties produced in city i at the asymmetric equilibrium when it is the larger city - over three times the number of varieties produced in the smaller city j. The lower panels show that a little over half of the varieties produced in each city are distributed to the other city. This share is independent of city size, as shown below in the proof of Lemma 2 in Appendix 1.6.

Figure 1.3 demonstrates Proposition 2: there are more varieties available in the larger city; differentiated product varieties are sold at a lower average price in the larger city, and these variety and price effects yield a lower differentiated product price index in the larger city. As noted in Appendix ?? below, each of these results depends on the fact that

more varieties are produced in the larger location.

Finally, note that the minimum housing requirement parameter, β, yields the non-linearity in the relative utility across the two cities and is, therefore, responsible for the existence of asymmetric equilibria. To see this, let us first look at the long-run equilibria where $\beta = 0$. Under this assumption, the indirect utility in city i relative to city j can be expressed as the sum of just two terms:

$$\hat{V}(\lambda, \Theta(\beta = 0)) = -\alpha\hat{P}(\lambda, \Theta) - (1 - \alpha)\hat{r}(\lambda, \Theta)$$

Equation (1.23), derived in the proof of Proposition 2 below, provides a simple expression for the relative differentiated product price index between the two cities when some, but not all, differentiated products are exported from each location:

$$\frac{P_i}{P_j} = \left(\frac{L_i}{L_j}\right)^{\frac{1}{1-\sigma}} \left(\frac{I - \beta r_i}{I - \beta r_j}\right)^{\frac{1}{1-\sigma}}$$

Equations (1.7) implies that the relative rent in city i to city j is equal to:

$$\frac{r_i}{r_j} = \frac{(1 - \alpha)IL_i(\bar{H} - \alpha\beta r_i)^{-1}}{(1 - \alpha)IL_j(\bar{H} - \alpha\beta r_j)^{-1}}$$

When when $\beta = 0$, both of the relative price index and relative rent terms above simplify such that:

$$\hat{P}(\lambda, \Theta) = \log\left(\frac{P_i}{P_j}\right) = \frac{1}{1-\sigma}\log\left(\frac{L_i}{L_j}\right) \text{ and } \hat{r}(\lambda, \Theta) = \log\left(\frac{r_i}{r_j}\right) = \log\left(\frac{L_i}{L_j}\right)$$

The relative indirect utility is, therefore, proportional to the log relative population:

$$\hat{V}(\lambda, \Theta(\beta = 0)) = -\left(\frac{1 - (1 - \alpha)\sigma}{1 - \sigma}\right)\log\left(\frac{L_i}{L_j}\right) = -\left(\frac{1 - (1 - \alpha)\sigma}{1 - \sigma}\right)\log\left(\frac{\lambda}{1 - \lambda}\right) \tag{1.16}$$

Under these assumptions, there can never exist any asymmetric equilibria. The relative utility is either constantly increasing or decreasing in relative population and, therefore, can only intersect the horizontal axis once, at the symmetric equilibrium. This is demonstrated in Figure (1.4). The red dotted line shows that, for $\beta = 0$, there exists one unstable symmetric equilibrium.

When $\beta = 0$, this model closely resembles Helpman (1998) who finds three different sets of dynamics. The first two have a single symmetric long-run equilibria, either stable or unstable, and are equivalent to the symmetric equilibria when the coefficient in equation (1.16), $-(1 - (1 - \alpha)\sigma)/(1 - \sigma)$, is negative and positive, respectively. Helpman (1998) also discusses a third set of dynamics that arises for some intermediate level of transport costs when $(1 - \alpha)\sigma < 1$. This set of dynamics also permits stable asymmetric equilibria in which both locations are populated. This set of dynamics cannot arise in this model when $\beta = 0$, production occurs in both locations, and there is bilateral trade. This is due to the fact that we have allowed for the intermediate input in final goods production to be freely traded, which means that in this case the small city can produce and export only the intermediate good while the large city will be the sole source of the final good.

Figure 1.4 shows that, as you increase β, the pattern of relative utility changes in two respects. First, the slope of relative utility becomes lower at each value of λ. Second, the slope is now non-linear and decreasing as the distribution of population becomes more asymmetric away from $\lambda = 0.5$. This is due to the fact that, for $\beta > 0$, rents increase non-linearly with city size. The minimum housing requirement is an inelastic component of individual housing demand. As the population increases in a location, the inelastic component of aggregate city-wide housing demand increases relative to the inelastic supply of housing in the city. This causes rents to increase more than proportionately with relative city size for any positive minimum housing requirement, and increasingly so as the minimum housing requirement increases. This non-linearity yields asymmetric equilibria in which consumers have not fully agglomerated.

With a positive minimum housing requirement, rents act as a congestion cost for both consumers and final goods producers. Consumers in the smaller city pay lower rents and, therefore, have more disposable income to spend on differentiated products after they have covered their minimum housing requirement. The disposable income of each small city resident is growing even as its total population decreases. This stems the flow of manufacturing production out of the small city into the large city and increases the range of λ for which the small city can maintain a large enough total expenditure to warrant some manufacturing production to occurring there. The minimum housing requirement, therefore, also plays a role in ensuring that firms do not fully agglomerate, such that

differentiated products are produced in both locations and traded in both directions.

□

1.5 Prices and Variety in Equilibrium

In this section, I examine the short-run equilibria for the above system of cities to study how and why differentiated product price index varies with city size.

The aggregate price level of differentiated products can be lower in the larger location either because there are more varieties available there, because these varieties are available at lower prices, or for both of these reasons. To see this, note that the relative price index for differentiated goods between two cities can be written as

$$\frac{P_i}{P_j} \ = \ \left(\frac{s_{ii} + (1 - s_{ii})\tau^{1-\sigma}}{s_{jj} + (1 - s_{jj})\tau^{1-\sigma}} \right)^{\frac{1}{1-\sigma}} \left(\frac{\tilde{n}_i}{\tilde{n}_j} \right)^{\frac{1}{1-\sigma}} \tag{1.17}$$

where $\tilde{n}_i = n_{ii} + n_{ji}$ is the total number of varieties available in city i and $s_{ii} = n_{ii}/(n_{ii} + n_{ji})$ is the share of those varieties that are produced locally.[8]

Equation (1.17) provides a decomposition of the price of differentiated goods into two terms: a variety effect capturing the fact that consumers prefer cities with more goods available and the Krugman price effect capturing the fact that the prices of locally-produced varieties are lower than imported ones. When all goods are traded there is an equal number of varieties available in each location, so $\tilde{n}_i = \tilde{n}_j = n_{ii} + n_{jj}$ and the second term of the relative price index defined in (1.17) drops out. In this case, the relative price index depends only on the shares of local products in consumption via the first "price effect" term. If no goods are traded, $s_{ii} = s_{jj} = 1$, so the first term of drops out and the relative price index depends only on the relative number of varieties available (and, in this case, produced) in each city via the second "variety effect" term. What is new in this model, relative to Krugman (1991), is that it is possible that some, but not all goods are traded, and the relative price index is decreasing in both the relative number of varieties available in each city and the shares of these varieties that are produced locally. That is, the price and variety effects both work simultaneously to lower the price index in the

---

[8]This decomposition is derived in Appendix 1.A.

larger location.

Before showing that these price, variety, and price index results, I first present two intermediate results on which these results depend. In what follows, I show that, under certain conditions, consumers in the larger city face lower average differentiated product prices, more differentiated product varieties, and, therefore, a lower differentiated product price index. These conditions are that differentiated products are produced in both cities and that some, but not all, of these products are traded in both directions. These three main results rely on the fact that, under these same conditions, an identical share of differentiated product firms export from each city and more products will be produced in the larger city. I prove each of these intermediate results in turn below.

Lemma 1. *When differentiated products are produced in both cities and some, but not all, of these products are traded in both directions, an identical share of differentiated product firms export their products from each city to the other city and, furthermore, these firms earn identical ex-post export profits.*

Proof. The most interesting case and empirically-relevant case, for the purposes of this paper, is that in which some, but not all, firms export from each city such that $f_{ij}^* = \pi_{ij}$ and $f_{ji}^* = \pi_{ji}$. We refer to these short-run equilibria as "interior" equilibria. In this case, the zero profit cutoff conditions implied by (1.14) define the share of firms exporting from each city in terms of ex-post export profits:

$$\frac{n_{ij}}{n_i} = \frac{\pi_{ij} - \underline{f_X}}{\overline{f_X} - \underline{f_X}} \quad \text{and} \quad \frac{n_{ji}}{n_j} = \frac{\pi_{ji} - \underline{f_X}}{\overline{f_X} - \underline{f_X}} \tag{1.18}$$

To solve for these export shares, we substitute these interior zero profit conditions, (1.18), into each of the free entry conditions implied by (1.15), using the fact that, by (1.12), and (1.13), $\pi_{ij} = \pi_{jj}\tau^{1-\sigma}$ and $\pi_{ji} = \pi_{ii}\tau^{1-\sigma}$[9]. This substitution yields two symmetric

---

[9]By definition, we have that $c_{ij} = c_{ji} = \tau$ and $c_{ii} = c_{jj} = 1$. Substituting these values into (1.12) and (1.13) yields

$$\pi_{ij} = \frac{\alpha(I - \beta r_j)L_j\tau^{1-\sigma}}{\sigma\left(n_{jj} + n_{ij}\tau^{1-\sigma}\right)} = \pi_{jj}\tau^{1-\sigma} \text{ and } \pi_{ji} = \frac{\alpha(I - \beta r_i)L_i\tau^{1-\sigma}}{\sigma\left(n_{ii} + n_{ji}\tau^{1-\sigma}\right)} = \pi_{ii}\tau^{1-\sigma}.$$

equations in the two unknown export shares, m_i and m_j:

$$f = \tau^{\sigma-1}\left[\frac{n_{ji}}{n_j}(\overline{f_X} - \underline{f_X}) + \underline{f_X}\right] + \frac{1}{2}\left(\frac{n_{ij}}{n_i}\right)^2(\overline{f_X} - \underline{f_X})$$

$$f = \tau^{\sigma-1}\left[\frac{n_{ij}}{n_i}(\overline{f_X} - \underline{f_X}) + \underline{f_X}\right] + \frac{1}{2}\left(\frac{n_{ji}}{n_j}\right)^2(\overline{f_X} - \underline{f_X})$$

This indicates that export shares depend only on the fixed cost of entry, f, the bounds of the fixed export cost distribution, $[\underline{f_X}, \overline{f_X}]$, and the iceberg transport cost, τ. More importantly, however, it indicates that the share of firms exporting is equal across the two cities. By the interior zero profit conditions, (1.18), this implies that export profits are also equal

$$\pi_X = \pi_{ij} = \pi_{ji} = m(\overline{f_X} - \underline{f_X}) + \underline{f_X} \tag{1.19}$$

where π_X denotes firm-level export profits and $m = \frac{n_{ij}}{n_i} = \frac{n_{ji}}{n_j}$ denotes the share of firms in each city that export. $\square$

Proposition. *When differentiated products are produced in both cities and some, but not all, of these products are traded in both directions:*

  1. *A higher number of differentiated product varieties are available in the larger city;*

  2. *Differentiated products are sold at a lower price, on average, in the larger city; and*

  3. *The differentiated product price index, P_i, is lower in the larger city.*

  Before presenting the proof of each part of this proposition, we first prove two intermediate results that will be useful below.

Lemma 2. *When differentiated products are produced in both cities and some, but not all, of these products are traded in both directions, more differentiated products are produced in the larger city.*

Proof. To show that more products are produced in the larger city, we first derive an expression for ex-post export profits using equations (1.11), (1.12), and (1.13):

$$\pi_X = \pi_{ij} = \frac{\alpha(I - \beta r_j)L_j\tau^{1-\sigma}}{(\sigma-1)(n_{jj} + n_{ij}\tau^{1-\sigma})}$$

$$\pi_X = \pi_{ji} = \frac{\alpha(I - \beta r_i)L_i\tau^{1-\sigma}}{(\sigma - 1)\left(n_{ii} + n_{ji}\tau^{1-\sigma}\right)}$$

We have assumed that some, but not all firms, export, so we can substitute these expressions for ex-post export profits into the interior zero profit cutoff conditions defined in equation (1.18), and assume that the same share of firms export from each location and substitute for $n_{ji} = mn_{jj}$ and $n_{ij} = mn_{ii}$. These substitutions yield two equations that jointly define the number of products produced in each location, as a function of parameter values and the size of and rents in each location:

$$n_{jj} + n_{ii}m\tau^{1-\sigma} = \alpha(I - \beta r_j)L_j\tau^{1-\sigma}/((\sigma - 1)\pi_X)$$

$$n_{ii} + n_{jj}m\tau^{1-\sigma} = \alpha(I - \beta r_i)L_i\tau^{1-\sigma}/((\sigma - 1)\pi_X)$$

The solution to this system of equations is as follows:

$$n_{ii} = \frac{\alpha\tau^{1-\sigma}\left((I - \beta r_i)L_i - (I - \beta r_j)L_j m\tau^{1-\sigma}\right)}{(\sigma - 1)\pi_X\left(1 - (m\tau^{1-\sigma})^2\right)}$$

$$n_{jj} = \frac{\alpha\tau^{1-\sigma}\left((I - \beta r_j)L_j - (I - \beta r_i)L_i m\tau^{1-\sigma}\right)}{(\sigma - 1)\pi_X\left(1 - (m\tau^{1-\sigma})^2\right)}$$

The number of varieties produced in city i relative to city j is, therefore,

$$\frac{n_{ii}}{n_{jj}} = \frac{\left((I - \beta r_i)L_i - (I - \beta r_j)L_j m\tau^{1-\sigma}\right)}{\left((I - \beta r_j)L_j - (I - \beta r_i)L_i m\tau^{1-\sigma}\right)}$$

Note that the relative number of varieties produced depends on three factors: (i) the relative population between the two cities, which we will denote using $\Lambda \equiv L_i/L_j$, (ii) the relative level of disposable income consumer's enjoy in city i relative to city j, $D \equiv (I - \beta r_i)/(I - \beta r_j)$, and (iii) the extent of the trade frictions between the two cities, $m\tau^{1-\sigma} = \tau^{1-\sigma}(\pi_X - \underline{f_X})/(\overline{f_X} - \underline{f_X})$. The relative number of varieties between the two cities, denoted by $N \equiv n_{ii}/n_{jj}$, is equal to

$$N = \frac{\left(\Lambda D - m\tau^{1-\sigma}\right)}{(1 - \Lambda D m\tau^{1-\sigma})}$$

In the symmetric equilibrium, populations and rents are equal across the cities so $\Lambda = 1$ and $D = 1$, and, therefore, $N = 1$ such that the same number of products is produced in each location. Outside the symmetric equilibrium, we know that however, relative populations and rents play a role in determining relative production variety. Our aim here is to show that, outside the symmetric equilibrium, there are more varieties produced in the larger city, such that $N > 1$ where $\Lambda > 1$ and $N < 1$ where $\Lambda < 1$. This will be the case if N is increasing in Λ, or

$$\frac{\partial N}{\partial \Lambda} \;=\; \frac{D(1 - (m\tau^{1-\sigma})^2)}{(1 - \Lambda D m\tau^{1-\sigma})^2} > 0$$

We have assumed that rents and the minimum housing requirement are low enough in both cities such that consumers have positive disposable incomes and $D > 0$. Therefore, this inequality holds as long as $(1 - (m\tau^{1-\sigma})^2) > 0$, which is true for all feasible transport costs, $\tau \geq 1$, by the assumption that some, but not all firms export, or $m < 1$.   □

It follows directly from Lemmas 1 and 2 that, under the same conditions, more varieties will be available to consumers in the large city than in the small city:

Proposition 2. *When differentiated products are produced in both cities and some, but not all, of these products are traded in both directions, a higher number of differentiated product varieties are available in the larger city.*

Proof. Here we need to show that:

$$n_{ii} + n_{ji} > n_{jj} + n_{ij} \tag{1.20}$$

for $L_i > L_j$.

Lemma 1 implies that $n_{ji} = mn_{jj}$ and $n_{ij} = mn_{ii}$, so the above inequality can be expressed as:

$$n_{ii} + mn_{jj} \;>\; n_{jj} + mn_{ii}$$
$$n_{ii}(1 - m) \;>\; n_{jj}(1 - m)$$
$$n_{ii} \;>\; n_{jj}$$

We have shown that this inequality holds for $L_i > L_j$ in Lemma (2). Therefore, equation 1.20 holds when $L_i > L_j$, *i.e.* there are more differentiated product varieties available in the larger of the two cities. $\qquad\square$

Krugman (1991) showed that, when trade costs are positive, but non-infinite, the average price of differentiated products is lower in the larger city and, since all goods are traded, this implies that the differentiated product price index is also lower there. Proposition 3 and 4 below show that these results are robust in a model in which some but not all goods are traded.

Proposition 3. *When differentiated products are produced in both cities and some, but not all, of these products are traded in both directions, differentiated products are sold at a lower price, on average, in the larger city.*

Proof. Equation (1.10) implies that the average price of differentiated products in city i, $\tilde{p}_i$ is equal to:

$$\tilde{p}_i = \frac{n_{ii} + n_{ji}\tau}{\rho(n_{ii} + n_{ji})}$$

We need to show that $\tilde{p}_i < \tilde{p}_j$ or alternatively that:

$$\frac{n_{ii} + n_{ji}\tau}{n_{ii} + n_{ji}} < \frac{n_{jj} + n_{ij}\tau}{n_{jj} + n_{ij}} \tag{1.21}$$

for $L_i > L_j$.

Lemma 1 implies that $n_{ji} = mn_{jj}$ and $n_{ij} = mn_{ii}$, so the above inequality can be simplified as follows:

$$
\begin{aligned}
\frac{n_{ii} + mn_{jj}\tau}{n_{ii} + mn_{jj}} &< \frac{n_{jj} + mn_{ii}\tau}{n_{jj} + mn_{ii}} \\
(n_{ii} + mn_{jj}\tau)(n_{jj} + mn_{ii}) &< (n_{jj} + mn_{ii}\tau)(n_{ii} + mn_{jj}) \\
n_{ii}n_{jj} + mn_{ii}^2 + mn_{jj}^2\tau + m^2 n_{ii}n_{jj}\tau &< n_{ii}n_{jj} + mn_{jj}^2 + mn_{ii}^2\tau + m^2 n_{ii}n_{jj}\tau \\
n_{jj}^2(1-\tau) &< n_{ii}^2(1-\tau) \\
n_{jj} &< n_{ii}
\end{aligned}
$$

We have shown that this inequality holds for $L_i > L_j$ in Lemma (2). Therefore, equation 1.21 holds when $L_i > L_j$, *i.e.* differentiated product varieties are sold at a lower average

price in the larger of the two cities.   □

The final result below states that the differences in variety and average prices outlined above combine to yield the key agglomerative force in the model, the lower differentiated product price index:

Proposition 4. *When differentiated products are produced in both cities and some, but not all, of these products are traded in both directions, the differentiated product price index, P_i, is lower in the larger city.*

Proof. Here we need to show that $P_i < P_j$ or, alternatively, that

$$\frac{P_i}{P_j} < 1$$

when when $L_i > L_j$.

We first use equations (1.10) to (1.13) to express ex-post export profits in terms of city-specific price indices as follows:

$$\pi_{ij} = \frac{\alpha(I - \beta r_j)L_j \tau^{1-\sigma}}{\sigma P_j^{1-\sigma}} \text{ and } \pi_{ji} = \frac{\alpha(I - \beta r_i)L_i \tau^{1-\sigma}}{\sigma P_i^{1-\sigma}} \tag{1.22}$$

Since these profits are equal by Lemma (1) under the conditions outlined in this proposition, we can use equation (1.22) to write the relative differentiated product price index between the two cities as the product of two terms:

$$\frac{P_i}{P_j} = \left(\frac{L_i}{L_j}\right)^{\frac{1}{1-\sigma}} \left(\frac{I - \beta r_i}{I - \beta r_j}\right)^{\frac{1}{1-\sigma}} \tag{1.23}$$

The first term directly relates the relative price index to relative city size. Since $\sigma > 1$, this term tends to make the differentiated price index lower in the larger city. The second term relates the relative price index to city size through city rents. Specifically, this term indicates that the relative price index is decreasing in the amount of income that consumers have for discretionary spending, after they have purchased the minimum housing requirement, β, in city i relative to city j.[10]

---

[10]Appendix Section 1 discussed why the minimum housing requirement is necessary for the empirically relevant heterogeneous cities equilibrium to exist .

The higher the rents in a city, the more expensive the minimum housing requirement and the lower the amount of income consumers have for discretionary spending. Since supply of housing is fixed in each location, rents are increasing in city size.[11] This indicates that the second term in the relative price index equation above puts upward pressure on the relative price index in the larger city. To see this, we take the difference between discretionary spending in city i relative to city j and substitute for equilibrium rents using equation (1.7) and then for income using equation (1.9):

$$
\begin{aligned}
(I - \beta r_i) - (I - \beta r_j) &= \beta(r_j - r_i) \\
&= (1 - \alpha)\beta I \left[\frac{L_j}{\bar{H} - \alpha\beta L_j} - \frac{L_i}{\bar{H} - \alpha\beta L_i} \right] \\
&= (1 - \alpha)\beta I \left[\frac{L_j(\bar{H} - \alpha\beta L_i) - L_i(\bar{H} - \alpha\beta L_j)}{(\bar{H} - \alpha\beta L_i)(\bar{H} - \alpha\beta L_j)} \right] \\
&= \frac{(1 - \alpha)\beta I \bar{H}(L_j - L_i)}{(\bar{H} - \alpha\beta L_i)(\bar{H} - \alpha\beta L_j)}
\end{aligned}
$$

The denominator of this term is positive as long as consumers in each city buy their minimum housing requirement, such that $\bar{H} > \beta L_i > \alpha\beta L_i$ and $\bar{H} > \beta L_j > \alpha\beta L_j$. The numerator is negative for $L_i > L_j$. Consumers' discretionary spending is, therefore, lower in the larger city. This tends to make the relative price index higher in this location.

To determine whether the price index is lower in the larger location, despite this discretionary spending effect, we now look at the first or second terms in the price index equation (1.23) together. Let λ equal the share of total population living in city i, such that $L_i = \lambda L$ and $L_j = (1 - \lambda)L$, and denote the per capita housing endowment as $\bar{h} = \bar{H}/L$. Using this notation and re-arranging terms, we can rewrite the relative price index as:

$$
\frac{P_i}{P_j} = \left[\frac{\lambda \left(\bar{h} - \alpha\beta(1 - \lambda) \right) \left(\bar{h} - \beta\lambda \right)}{(1 - \lambda) \left(\bar{h} - \beta(1 - \lambda) \right) \left(\bar{h} - \alpha\beta\lambda \right)} \right]^{\frac{1}{1 - \sigma}}
$$

In order to show that the price index is lower in the larger city, we need to show that the

---

[11]To see this mathematically, note that the derivative of rents, defined in equation (1.7), with respect to city size is

$$
\frac{\partial r_i}{\partial L_i} = \frac{(1 - \alpha)I\bar{H}}{(\bar{H} - \alpha\beta L_i)^2}
$$

Since the Cobb-Douglas expenditure share on differentiated products, α, is between zero and one, this derivative is positive.

relative price index is less than one for $\lambda > 0.5$ or, alternatively, that the numerator of the term in the square brackets is greater than the denominator for $\lambda > 0.5$:

$$
\begin{aligned}
P_i^{1-\sigma} - P_j^{1-\sigma} &= \lambda \left(\bar{h} - \alpha\beta(1-\lambda)\right)\left(\bar{h} - \beta\lambda\right) - (1-\lambda)\left(\bar{h} - \beta(1-\lambda)\right)\left(\bar{h} - \alpha\beta\lambda\right) \\
&= \beta\left(\alpha(1-\lambda)\lambda + (1-\lambda)^2 - \alpha\lambda(1-\lambda) + \lambda^2\right) \\
&= \beta\left((1-\lambda)^2 - \lambda^2\right) \\
&> 0
\end{aligned}
$$

$\square$

1.6 Conclusion

This paper has presented a simple NEG model that generates asymmetric equilibria where residents of large cities benefit not only from the lower prices, but also from the greater variety available to them relative to residents of small cities. The basic intuition is that agglomeration will result in lower goods prices but also yield congestion costs, with the added twist along the lines of Melitz (2003) that not all goods will be consumed everywhere. Here, as in Helpman (1998)'s variant of the seminal Krugman (1991) framework, congestion costs are due to the fact that the price of a fixed supply of land, or housing, increases with a city's population. Just as Krugman (1991) and Helpman (1998) showed that there can be asymmetric urban equilibria characterized by larger cities having lower prices, this paper showed that adding heterogeneous fixed costs of exporting to these models yields asymmetric equilibria in which larger cities have *both* lower prices and more varieties of traded goods.

1.A Derivation of Price Index Decomposition

Proposition. *The relative price index for differentiated goods between two cities can be written as:*

$$\frac{P_i}{P_j} = \left(\frac{\tilde{n}_i}{\tilde{n}_j}\right)^{\frac{1}{1-\sigma}} \left(\frac{s_{ii} + (1-s_{ii})\tau^{1-\sigma}}{s_{jj} + (1-s_{jj})\tau^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} \tag{1.24}$$

where $\tilde{n}_i = n_{ii} + n_{ji}$ is the total number of varieties available in city i and $s_{ii} = n_{ii}/(n_{ii} + n_{ji})$ is the share of those varieties that are produced locally.

Proof. The price index for a city $j \in \mathbb{I}$ is defined in equation (1.11) rewritten below:

$$P_j = \left(\sum_{i \in \mathbb{I}} n_{ij} p_{ij}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = \frac{1}{\rho}\left(\sum_{i \in \mathbb{I}} n_{ij} c_{ij}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$

Using the definition of marginal costs, c_{ij}, from equation (1.10) and defining the set of two cities to be $\mathbb{I} = \{i, j\}$, we can re-write the price index as:

$$
\begin{aligned}
P_j &= \frac{1}{\rho}\left(n_{jj} + n_{ij}\tau^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\
&= \frac{1}{\rho}(n_{jj} + n_{ij})^{\frac{1}{1-\sigma}}\left(\frac{n_{jj}}{n_{jj} + n_{ij}} + \frac{n_{ij}\tau^{1-\sigma}}{n_{jj} + n_{ij}}\right)^{\frac{1}{1-\sigma}} \\
&= \frac{1}{\rho}(\tilde{n}_j)^{\frac{1}{1-\sigma}}\left(s_{jj} + (1-s_{jj})\tau^{1-\sigma}\right)^{\frac{1}{1-\sigma}}
\end{aligned}
$$

where we use $\tilde{n}_j = n_{jj} + n_{ij}$ to denote the number of varieties available in city j and $s_{jj} = n_{jj}/(n_{jj} + n_{ij})$ to denote the share of those varieties that are produced locally in city j. The relative price index between two cities i and j can, therefore be written as:

$$\frac{P_i}{P_j} = \left(\frac{\tilde{n}_i}{\tilde{n}_j}\right)^{\frac{1}{1-\sigma}} \left(\frac{s_{ii} + (1-s_{ii})\tau^{1-\sigma}}{s_{jj} + (1-s_{jj})\tau^{1-\sigma}}\right)^{\frac{1}{1-\sigma}}$$

$\square$

Figure 1.1: Utility in City i Relative to City j $(\log(U_i/U_j))$ vs. Proportion of Population in City i (λ)
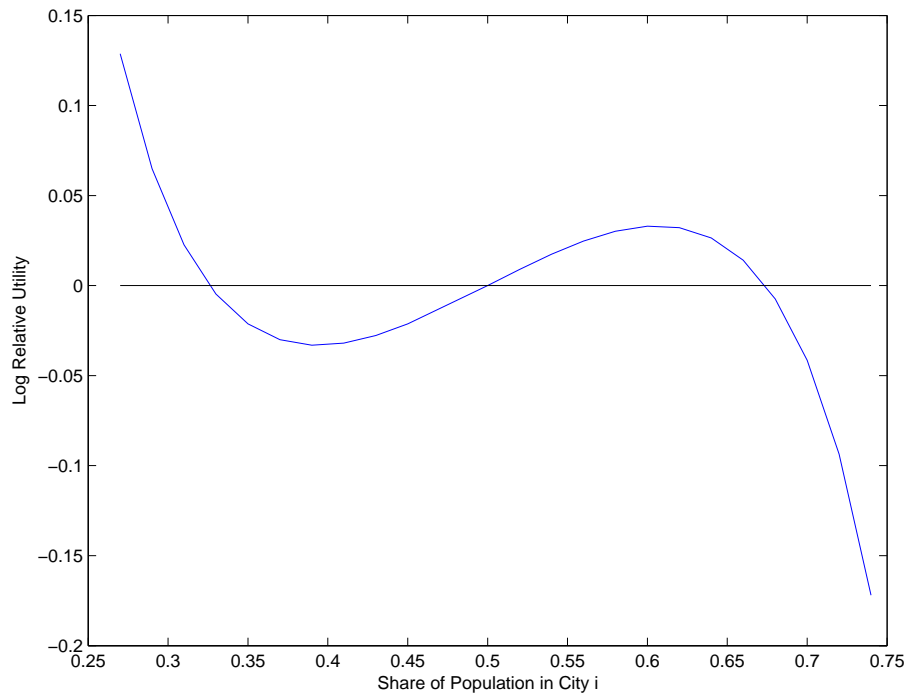
Figure 1.2: Number of Varieties Produced in Each City and Share Exported vs. Proportion of Population in City i (λ)
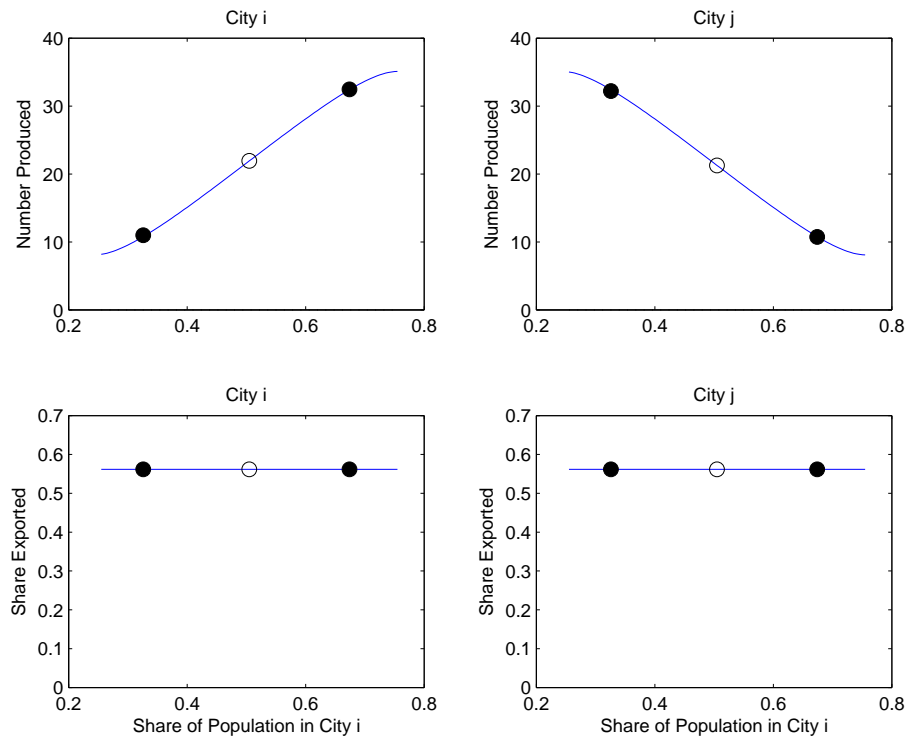
Figure 1.3: Variety, Prices, and Price Indices in City i Relative to City j $(\log(U_i/U_j))$ vs. Proportion of Population in City i (λ)
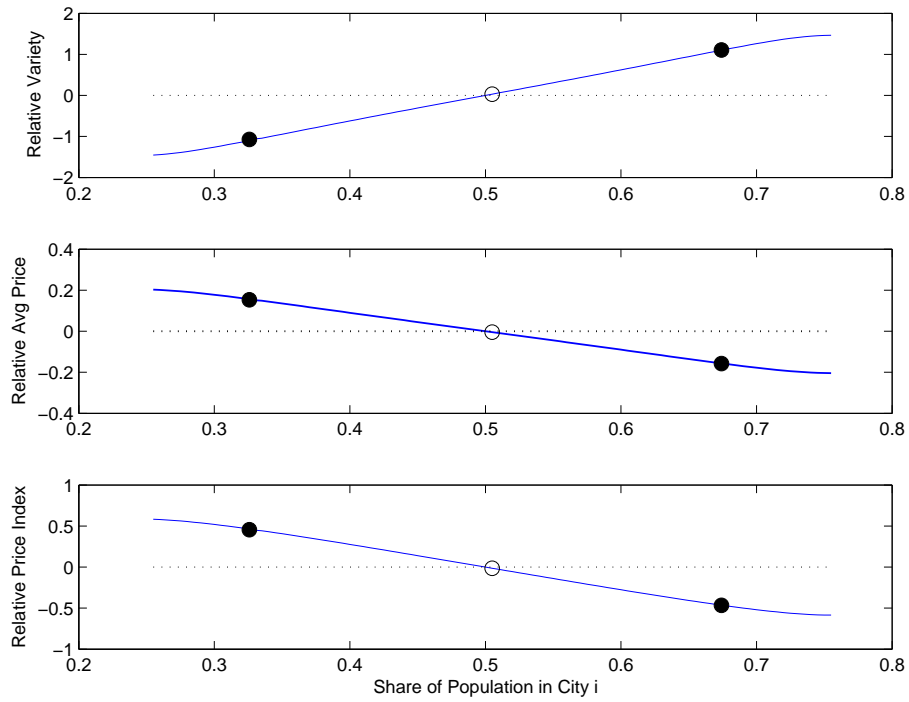
Figure 1.4: Utility in City i Relative to City j $(\log(U_i/U_j))$ vs. Proportion of Population in City i (λ) for Different Minimum Housing Requirements (β)
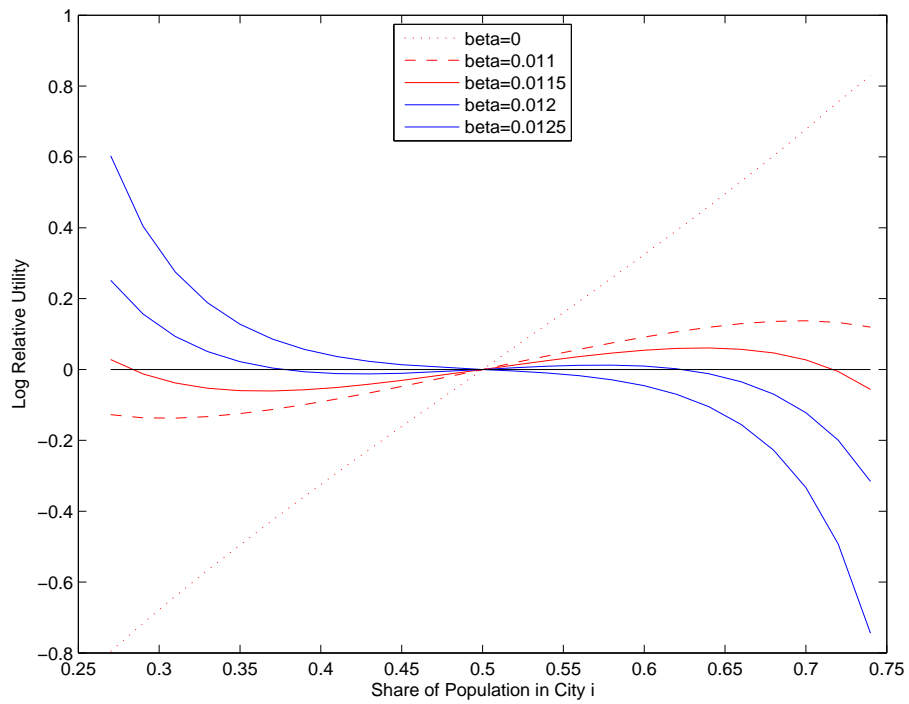
Chapter 2

Is New Economic Geography Right? Evidence from Price Data

Jessie Handbury and David E. Weinstein[1]

2.1 Introduction

In awarding Paul R. Krugman the Nobel Prize in economics for his work for international trade and economic geography, the committee characterized his contribution to economic geography as follows: "The new economic geography initiated by Krugman broke with ... tradition by assuming *internal* economies of scale and imperfect competition. Agglomeration is then driven by *pecuniary externalities* mediated through market prices as a large market allows greater product variety and lower costs [emphasis in original]."[2] Despite this award, there have been no tests of whether larger markets actually do have greater product variety and whether differences in the number of available varieties are sufficient to lower costs. This paper is the first to fill this gap by testing empirically whether large markets are actually characterized by lower prices of traded goods, and more importantly, whether this, in conjunction with greater product variety, lowers costs for consumers. In other words, is the fundamental mechanism underlying new economic geography models correct?

Over the last two decades, there has been a burgeoning literature that has aimed to test the implications of New Economic Geography (NEG) models (see the recent excellent surveys by Brakman, Garretsen, and van Marrewijk (2009); Fujita and Mori (2005); Combes, Mayer, and Thisse (2008)) as well as the costs of remoteness such as Redding and Sturm (2008). Similarly, Combes, Duranton, Gobillon, Puga, and Roux (2009) have developed models and empirical tests that underscore the importance of agglomeration economies relative to firm selection. These tests have examined predictions of the theory – e.g., home market effects, multiple equilibria, patterns of spatial agglomeration, and wage distribution – but they left untouched the question of whether the precise agglomeration mechanism Krugman postulated to be important was at work.

This paper deviates from prior work by using highly disaggregated barcode data covering hundreds of thousands of goods purchased by 33,000 households in 49 cities in the US to measure product variety and the variety-adjusted costs faced by consumers in different locations. This permits the precise estimation of the pecuniary externalities hypothesized by Krugman.

NEG models generate externalities from a very simple structure. Trade costs are positive between cities, so locally-produced goods must be cheaper in the local market than they are elsewhere. Larger cities will produce more varieties than smaller ones, so consumers in these locations will purchase more locally-produced goods and, therefore, face a lower average price. When not all products are traded, consumers will also have access to more varieties of goods in larger cities. These two forces combine to generate a lower price index in larger cities.[3] This lower price index induces people to move into cities enhancing the scale economy.

The notion that prices are lower in larger cities has not been supported by data examined in previous work. In a seminal paper estimating a general equilibrium model of city size, Roback (1982) identified a cost-of-living premium using U.S. housing prices and, more recently, DuMond, Hirsch, and Macpherson (1999) confirmed this result using aggregate U.S. cost-of-living indices. Similarly, Tabuchi (2001) finds that aggregate consumer price

---

[3]In the original Krugman (1991) paper more varieties are produced in a large market but with non-infinite trade costs all varieties are available to consumers in both large and small markets. We will specifically refer to the fact that more varieties are available to consumers in large markets as "the variety effect." The previous chapter shows that this can arise in Helpman (1998) 's variant of the Krugman (1991) model when there is a heterogeneous fixed cost associated with trading goods between cities.

indices, land values, and housing costs are all higher in larger Japanese cities. One reason that these studies have not been deemed fatal for the theory is that it is easy to modify NEG models to generate higher housing prices in cities, as demonstrated in Helpman (1998) and in the theory section below, or higher non-traded goods prices, as in Suedekum (2006).

Prior empirical studies also suffer from biases because they do not compare the prices of identical goods. Broda, Liebtag, and Weinstein (2009), for example, find that wealthier households purchase substantially more expensive varieties of the same narrowly defined good – e.g., milk, eggs, water – than poorer households, even in the same store. Similarly, Broda and Weinstein (2008) find that most of the variation in average prices paid by wealthy and poor households for goods like milk, eggs, and water is due to differences in the varieties consumed not the prices of the underlying varieties. The remaining variation is due to differences in shopping behavior; wealthy households are more likely to shop at high-amenity stores and respond less to sales. To the extent that stores in wealthier neighborhoods offer more amenities and different ranges of products to cater more to local clientele, this is likely to bias studies based on non-identical samples of goods purchased in non-identical stores to find higher prices in wealthier locations. The Bureau of Labor Statistics accounts for these effects by computing the Consumer Price Index (CPI) using price changes of a precisely-defined bundle of goods *within* a given store reported by a field agent following instructions to control for shopping differences across consumers. Similarly, in this paper, we compare prices across cities using indices of the prices of identical products sold in the same store chain by the same type of shopper.

Our paper makes a number of contributions. Consistent with earlier analyses, we show that a simple price index for *identical goods* rises with city size. However, we also demonstrate that this result is not robust to controls for the household making the purchase (*i.e.,* shopping characteristics) and the type of store in which the purchase is made (*i.e.,* store amenities). In other words, most of the reason prices of identical goods appear to be more expensive in large cities is that residents of those cities tend to buy more expensive varieties and they purchase them more frequently in "nice" stores that are expensive in all locations. We find that prices for the same good purchased are actually lower in larger cities once we control for these forces. This is the first empirical confirmation of Krugman's

conjecture about how the prices of traded goods vary with city size. Moreover, if we control for the fact that prices in larger cities embody land prices, we find that prices net of land costs fall even more sharply with city size. In the strictest sense, the results on the prices of traded goods available in more than one city are a precise structural estimation of the Krugman (1991) agglomeration force, and one need not think about what the prices are for non-traded goods.

However, our study also is the first to document that the variety of traded goods available for consumption is substantially higher in larger cities and, as we show theoretically, this result is very much in line with the NEG approach. The elasticity of the number of products with respect to city size is a whopping 0.2-0.3 (depending on the specification) which means that there is enormous variation in the number of available varieties across cities. For example, we estimate residents of New York (population 9.3 million) can choose between 97,000 different types of groceries, whereas residents of Des Moines (population 456,000) only have access to 32,000 varieties. This greater availability of varieties in larger cities means that variety-adjusted costs are likely to be substantially lower in large cities, as NEG models predict, setting the stage for our econometric exercise seeking to quantify the importance of the availability of varieties for welfare.

To estimate these welfare effects, we construct a variety-adjusted exact price index for each city in our sample. Our results show that while identical goods are purchased at higher prices in larger cities, the variety effect offsets this price effect resulting in similar variety-adjusted costs across cities. Since the prices of identical goods are actually lower in larger cities when we control for purchaser characteristics and store amenities, the variety-adjusted costs are substantially lower for a consumer in a large city than for a consumer sharing the same characteristics shopping in the same stores in a small city. These effects are not small. Our estimates indicate that a household that moved from Des Moines to New York and purchased goods from the same type of stores in the two cities would realize a 10 percent drop in the overall cost of its grocery purchases. This suggests that the price effects hypothesized by Krugman are economically significant.

The rest of the paper is structured as follows. Section 2.2 describes the data we use and develops some stylized facts about purchasing behavior and variety availability that motivates our modeling choices. Section 2.3 develops the price index theory and

econometrics underlying our estimation strategy. We present our results in Section 2.4, and Section 2.5 concludes.

2.2 Data

The primary dataset that we use is taken from the Nielsen HomeScan database. This data was collected by Nielsen from a demographically representative sample of approximately 33,000 households in 52 markets across the U.S. in 2005. Households were provided with Universal Product Code (UPC) scanners to scan in every purchase they made including online purchases and regardless of whether purchases were made in a store with scanner technology.[4] We have the purchase records for grocery items, which include purchase quantity, price, date, store name, product "module" (*i.e.*, a detailed description of the type of good, *e.g.* diet carbonated beverages), as well as demographic information for each household making the purchase. For much of the analysis, we will be working with "brand-modules" that correspond to all the UPCs within a module that are marketed under a particular brand, *e.g.*, "Fanta-diet carbonated beverages." We will also use product "groups," a more aggregate product categorization provided by Nielsen. Each product "module" fits within a unique product "group" (*i.e.*, the "diet carbonated beverages" and the "regular carbonated beverages" product modules both fit within the "carbonated beverages" product group). Detailed descriptions of these data and the sampling methods used can be found in Broda and Weinstein (2010).

A major advantage of these data relative to the data used in previous studies comparing cost-of-living indices across cities is that we can compare prices of identical goods to directly test the prediction that traded goods are on average cheaper in agglomeration locations. We can further use the household level demographic data and store names to separate how much of the inter-city price differential arises from factors that are not considered in the NEG models, such as differences in demographics, shopping behavior, and/or store amenities.

Although the Nielsen dataset contains data for 52 markets, we classify cities at the

---

[4]In cases where panelists shop at stores without scanner technology, they report the price paid manually. Since errors can be made in this reporting process, we discard any purchase records for which the price paid was greater than twice or less than half the median price paid for the same UPC, approximately 250,000 out of 16 million observations.

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, we group them all as New York-Northern New Jersey-Long Island CMSA. There are two cases in which Nielsen groups two MSAs into one market. In these cases, we count the two MSAs as one city, using the sum of the population and manufacturing output and the population-weighted mean land value. We use population, income distribution, and racial and birthplace diversity data from the 2000 U.S. Census[5] and 2000 land values from Davis and Palumbo (2007). Matching the Nielsen data with the Davis and Palumbo data reduces the number of cities in our analysis to 37.

Before turning to the theory and estimation, we characterize two stylized facts of the raw data that motivate our approach:

Stylized Fact 1: Grocery Unit Value Indices Indicate Large Cities are More Expensive   A common method to compare price levels across cities within countries relies on unit value indices such as those published by the Council for Community and Economic Research (formerly the American Chamber of Commerce Research Association (ACCRA)). The ACCRA Cost of Living Index (COLI) is a unit value index calculated using prices for similar products collected in different cities. The product categories are narrowly defined, *e.g.* "price per pound of ground beef, lowest price, min 80% lean," "12 oz. plain regular potato chips," or a "two bedroom, unfurnished apartment, excluding all utilities except water, 1 or 2 baths, 950 sq. ft." These indices have often been used to argue that prices are higher in larger cities. While we will deal with a host of problems arising from comparing prices of similar, as opposed to identical, products later, it is important to establish that the result that prices are higher in larger cities is not just driven by the housing component of these indices but is also true when looking at the grocery component alone.

In order to establish this stylized fact, we obtained the ACCRA COLI data for 2005 and used it to build a food index based on their unit values and weights. We also matched every food category used in the ACCRA COLI index with the same category of UPCs in

---

[5]The Nielsen sample is demographically representative within each market.

the Nielsen data and built a food index based on the average unit values of these UPCs and the ACCRA weights. We refer to these two constructed indices as the ACCRA food index and Nielsen food index.

Table 2.1 presents cross-city correlations between these various unit value indices and city size. The first column demonstrates that it is not just the aggregate COLI index that is highly and positively correlated with city size, but grocery sub-indices are also highly correlated with city size. The correlation between the aggregate COLI and log city population is 0.68 overall and 0.55 for food items. Similarly, the correlation between log population and the Nielsen food index is 0.56 – almost identical to that for the ACCRA food index. If we look at the correlations between the different indices in columns 2 and 3, we see that that correlation coefficient tends to hover around 0.8, indicating that food prices tend to be high whenever aggregate prices are high. In other words, it is not just indices containing land unit values that indicate large cities are more expensive, grocery unit values are also higher.

Table 2.1: Correlation Coefficients

	Log(Population$_c$)	Nielsen Food Index	ACCRA Food Index	ACCRA Composite Index
Log(Population$_c$)	1			
Nielsen Food Index	0.56	1		
ACCRA Food Index	0.55	0.77	1	
ACCRA Composite Index	0.68	0.81	0.85	1

Stylized Fact 2: Consumers in Larger Cities Consume More Varieties One simple way to see if households in larger cities have access to more varieties is to count them. Because Nielsen tends to sample more households in larger cities, we need to make sure that the sample size is constant across cities. We therefore restrict ourselves to only looking at cities in which Nielsen sampled at least 450 households and compare the number of varieties purchased by a random sample of 450 households in each of these cities. We plot the aggregate number of *different* UPCs purchased by all of these households against the size of the city in which the households live in Figure 2.1. The results show a clear positive relationship between the variety of UPCs purchased in a city and the population of the city. This is certainly suggestive of the notion at the center of NEG models that

the number of varieties available in a location rises with number of inhabitants in that location. However, it leaves open the question of whether these differences are likely to matter for consumers; that is, does the city size variety effect drive down the consumer price index in larger cities relative to small? We now turn to assessing this question.

Figure 2.1
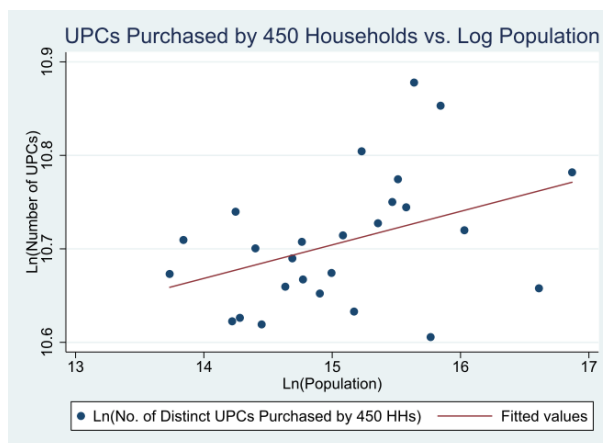
2.3   Empirical Strategy

In Section 2.4 we will examine the relationship between city size and consumer costs by answering three separate, but related questions: 1) Are traded goods prices lower in larger cities as predicted in Krugman (1991)?; 2) Are more varieties available in larger cities as predicted by the Krugman (1991) extension presented above?; and 3) What is the combined effect of the lower prices and greater variety on the aggregate price index faced in large, relative to small, cities; that is, what is the size of the price index effect that drives agglomeration in both Krugman (1991) and the model presented above?[6] To answer each of these question we need to measure traded goods prices, variety, and the effects of each component on the aggregate price index. We present the methods we use for these measurements here before presenting our empirical results below.

---

[6]Note that these questions directly test the main results of the previous chapter and, in doing so, also test the original Krugman (1991) predictions that traded goods prices are lower in larger cities and that the aggregate price index is lower in these locations, even without there being greater variety there.

2.3.1 Adjusting Prices for Amenities and Buyer Characteristics

Our first challenge in identifying the "price index effect" is to measure the prices of goods consumed in more than one city – *i.e.*, "common goods." One of the complicating factors associated with measuring these prices is that store amenities and shopping behavior vary systematically across locations. If, for example, bigger cities feature nicer stores, we might systematically bias our results against the NEG model. Similarly, wealthier households often pay more for the same goods in the same stores (presumably because the opportunity cost of time spent shopping is higher) and per capita income is increasing in city size. Finally, since urban population is positively correlated with land prices in most spatial models, the non-traded, land price component of grocery prices might drive a positive correlation between population and prices and one may want to strip this component out to evaluate the pure impact of population on tradable goods prices. This suggests that we should purge the data of these effects, but in the interests of understanding which effects are critical, we will neutralize these confounding factors one at a time.

We begin by letting P_{uzcrh} be the price of UPC u, purchased in zip code z, in store r, by household h in city c. We will refer P_{uzcrh} as the "unadjusted price" and define p_{uzcrh} as $\ln(P_{uzcrh})$. The simplest way to purge the price data of amenities and other effects that are not in the model is in a two step process, the first of which is to run a regression of the form:

$$p_{uzcrh} = \alpha_u + \gamma \ln(Pop_c) + \mathbf{Z}_h\beta + \delta_r + \phi_1 \ln(Income_z) + \phi_2 \ln(Land_c) + \varepsilon_{uzcrh} \quad (2.1)$$

where $\mathbf{Z}_h$ denotes a series of household characteristics; Pop_c denotes the population of the city in which the UPC was purchased; $Income_z$ is the per capita income in the zip code where the UPC was purchased; $Land_c$ is the price of land in the city where the UPC was purchased; and Greek symbols denote parameters to be estimated.

Equation (2.1) enables us to make a number of adjustments to the prices to control for various characteristics that can affect prices but are outside of the canonical Krugman model. We construct the log demographic-adjusted price as

$$\tilde{p}^D_{uzcrh} = p_{uzcrh} - \mathbf{Z}_h\widehat{\beta} \quad (2.2)$$

$\tilde{p}^D_{uzcrh}$ corresponds to log prices purged of effects of demographic characteristics of the buyers. This will be useful when we want to control for the fact that wealthier people tend to locate in larger cities. Crucially, we have included population in regression 2.1 so that even if household demographics, or any of the other controls, are correlated with city size, that will not affect the magnitude of our adjustment. We then purge prices of store effects in the log store-adjusted price defined as

$$\tilde{p}^{SD}_{uzcrh} = p_{urzcrh} - \mathbf{Z}_h\widehat{\beta} - \widehat{\delta}_r \tag{2.3}$$

For stores with more than $100,000 in sales, which includes national chains such as *Wal-Mart*, *Kroger*, and *CVS*, we estimate δ_r for each chain and purge prices of chain-specific store amenities. We therefore included close to 100 store dummies; the majority of these are located in 15 or more cities, and all in at least 2 cities. For smaller stores, however, we restrict δ_r to be the same for all stores of the same type, where type is defined in one of seven "channel-IDs": grocery, drug, mass merchandiser, supercenter, club, convenience, and other. Here, we are accounting for the fact that products might be distributed through different channels in larger cities than smaller cities, *e.g.* through convenience and grocery stores as opposed to super centers and club stores. $\tilde{P}^{SD}_{uzcrh} = \exp\left(\tilde{p}^{SD}_{uzcrh}\right)$ is the closest approximation we have to the BLS's approach of looking at prices within stores using a given shopping approach, and it forms the basis of our preferred specification.

However, for the sake of robustness, we also consider a number of other possibilities. Broda, Liebtag, and Weinstein (2009) found that stores in more expensive neighborhoods charge more for the same UPCs presumably because consumers may enjoy shopping in nicer neighborhoods and/or grocery-store workers in high income neighborhoods may offer more amenities. We account for this possibility defining the following income-adjusted price

$$\tilde{p}^{SDI}_{uzcrh} = p_{urzcrh} - \mathbf{Z}_h\widehat{\beta} - \widehat{\delta}_r - \widehat{\phi_1}\ln(Income_z) \tag{2.4}$$

These prices will be useful when we want to consider the shift in utility that would occur if a household with a particular set of demographic characteristics living in a particular type of neighborhood moved to a comparable neighborhood in another city.

Finally, since even the prices of tradables contain a non-traded component (and this is

something we could easily adjust our model to incorporate), we may also wish to consider prices that have also been purged of land values as an additional robustness check. We define land-value adjusted prices as

$$\tilde{p}_{uzcrh}^{SDILV} = p_{urzcrh} - \mathbf{Z}_h\widehat{\beta} - \widehat{\delta_r} - \widehat{\phi_1}\ln(Income_z) - \widehat{\phi_2}\ln(Land_c) \qquad (2.5)$$

These prices will give us a sense of how the prices of purely traded goods vary across cities before land rents are included in the retail prices.

2.3.2 The Price Level in Cities

We will adopt two approaches to the measurement of prices in cities. The first corresponds to a purely structural test of the pecuniary externality identified in Krugman (1991). Since all varieties are available for consumption everywhere, the proper way to aggregate the prices of tradable goods is the CES price index. This is something that we can easily compute for the set of goods available in any city. Checking whether this index falls with city size provides a direct test of the agglomeration mechanism in Krugman (1991).

However, we have also seen that a strong feature of the data is the greater availability of tradable varieties in larger cities. We have extended the basic model to reflect this phenomenon and now describe how we account for it in our empirical analysis. Feenstra (1994) developed the variety-adjusted price index for the CES utility function. Here, we will modify it so that it can be used with our data.[7] We begin by modeling the aggregate utility function as a nested CES utility function. To do this, we need to establish some notation. Let $g \in \{1, ..., G\}$ denote the set of product groups, which we define in the same way as Nielsen, *i.e.*, sectors like "Crackers", etc.. As in Broda and Weinstein (2010), the product group of "Crackers", would contain brand-modules that are the intersections of brands (*e.g.* "Nabisco-Premium") and modules within product groups (*e.g.* "Flaked Soda

---

[7]Given that the median number of UPCs purchased in a module by a single person household (conditional on purchasing anything in the module) is one, the data suggest that one should think of households as having heterogeneous ideal-type preferences, as opposed to the identical CES preferences that form the theoretical foundation for the variety-adjusted exact prices indices used in this paper. This discrepancy, however, is not a problem for our analysis if we think of consumers as having a logit demand system. In particular, Anderson, de Palma, and Thisse (1987) have demonstrated that a CES demand system can arise from the aggregation of ideal-type logit consumers. We will therefore follow Anderson, de Palma, and Thisse (1987) and use the CES structure to evaluate aggregate welfare even though we know that the discrete choice model is a better depiction of reality at the household level.

Crackers" or "Cheese Crackers"). Thus brand-modules would be categories like "Nabisco-Premium-Flaked Soda Crackers" and "Pepperidge Farm Goldfish-Cheese Crackers." We let $b_g \in \{1, ..., B_g\}$ denote the set of "brand-modules" within a product group.

We define U_{bgc} the set of all UPCs that have positive sales in city c in brand-module b, in product group g; U_{bg} the set of all UPCs that have positive sales *nationally* in brand-module b, in product group g; B_{gc} the set of all brands that have positive sales in city c in product group g; and B_g the set of all brands that have positive sales *nationally* in product group g.

It turns out that it is useful to measure the importance of a UPC's sales relative to several benchmarks. First, let s_{ubg} be the *national* expenditures on UPC u as a share of national expenditures on brand b in product group g, and let s_{ubgc} be a city's expenditures on UPC u as a share of the city's expenditures on brand b in product group g. We next define the national market share of city c's available UPCs in that particular brand-module (*i.e.*, $u \in U_{bgc}$) as

$$s_{bgc} \equiv \sum_{u \in U_{bgc}} s_{ubg}$$

Thus, s_{bgc}, which we call the "UPC-share," tells us the expenditure share of UPCs within a brand that are available in a city using *national* weights. In other words, if the UPCs that constitute a large share of a brand's national sales are available locally, then s_{bgc} will be large. It will also be useful to define an analogous measure computed using *local* weights:

$$\tilde{s}_{bgc} \equiv \sum_{u \in U_{bgc}} s_{ubgc}$$

Just as we defined shares of UPCs within brands, we also need to measure the importance of various brands. We do this in an analogous manner below;

$$s_{bgc} \equiv \sum_{b \in B_{gc}} s_{bg} \text{ and } \tilde{s}_{gc} \equiv \sum_{b \in B_{gc}} s_{bgc}$$

where s_{bg}, which we call the "brand-share," is the *national* market share of brand-module b in product group g, and s_{bgc} is the local market share of brand-module b in product group g. We can now modify Broda and Weinstein (2010)'s Proposition 1 as follows:

Proposition 5. *For $g \in G$, if $B_{gc} \neq \emptyset$ then the exact price index for the price of the set of goods G in city c relative to the nation as a whole that takes into account the differences in variety in the two locations is given by,*

$$EPI_c = \prod_{g \in G} \left[\prod_{u \in U_{gc}} CEPI_{gc} \left(s_{gc}\right)^{\frac{1}{1-\sigma_g^a}} \prod_{b \in B_g} \left(s_{bgc}\right)^{\frac{w_{bgc}}{1-\sigma_g^w}} \right]^{w_{gc}}$$

where w_{bgc} and w_{gc} are log-ideal CES [Sato (1976)](#) and [Vartia (1976)](#) weights defined as follows:

$$w_{gbc} = \frac{\frac{\tilde{s}_{bgc} - s_{bg}}{\ln \tilde{s}_{bgc} - \ln s_{bg}}}{\sum\limits_{b \in B_g} \left(\frac{\tilde{s}_{bgc} - s_{bg}}{\ln \tilde{s}_{bgc} - \ln s_{bg}} \right)} \quad and \quad w_{gc} = \frac{\frac{\tilde{s}_{gc} - s_g}{\ln \tilde{s}_{gc} - \ln s_g}}{\sum\limits_{g \in G} \left(\frac{\tilde{s}_{gc} - s_g}{\ln \tilde{s}_{gc} - \ln s_g} \right)}$$

σ_g^a is the elasticity of substitution across brand-modules in product group g, σ_g^w is the elasticity of substitution among UPCs within a brand-module, s_g is the share of product group g in national expenditures, and $CEPI_{gc}$ is the conventional exact price index associated with the CES utility function for product group g and city c relative to the national sample.

This index has been used extensively in the literature (see Feenstra (2010) for a survey of its use), so we will only briefly discuss its properties here. While the $CEPI_{gc}$ measures the prices of products available in city c relative to their average price nationally, the second and third terms within the brackets adjust the exact price index to account for the fact that not all varieties, respectively defined as brands and UPCs, available nationally are available in city c.

Each variety adjustment is calculated using an elasticity of substitution parameter and the national expenditure share of unavailable varieties. One of the properties of this index is that the variety adjustments and, therefore, the price level will fall as more UPCs and brands are available in a city, and will fall faster when these additional varieties and brands are popular nationally. This property is due to the fact that the index gives more weight to the availability of non-common varieties that comprise a larger share of expenditure in the market where they are available, in our case, the national market. The elasticity of substitution parameter weights varieties by how differentiated they are. As a result, the availability of more differentiated varieties matters more than the availability of varieties

that are close substitutes to existing varieties. We obtain the elasticities of substitution computed for UPCs within a brand-module and across brand modules within a product group from Broda and Weinstein (2010).

2.3.3   Measuring Variety Availability in Cities

Measuring the brand-share, s_{gc}, and the UPC-share, s_{bgc}, provides a particular challenge because we do not observe all of the varieties available in each city. Our estimates of s_{bgc} and s_{gc} are likely to be affected by a downward bias on our estimates of the size of sets U_{bgc} and U_{gc}. The bias is likely to emerge from two sources. First, some cities have smaller samples of households than other cities, so there will be a natural tendency for the counts of UPCs purchased in each city to vary with the sample size in the city. Second, even if the sample sizes were identical in every city, the sample count of UPCs will be below the true count of UPCs because some goods will not be purchased by our sample of households but will be purchased by other households in the city.

Our approach to this problem is to break it into two components. We begin by developing two methodologies for estimating the number of UPCs and brand-modules available in a city, the size of the sets U_{bgc} and U_{gc} respectively, using purchase records for only a sample of households in the city. The first method is parametric and the second is non-parametric, but we will show that both methodologies yield very similar estimates for how the numbers of varieties and brands vary across cities, which is a critical component in understanding whether the extended NEG model is valid.

Unfortunately, there is not a theoretical underpinning for parametrically estimating s_{gc} and s_{bgc}. Nevertheless, these can be estimated using a non-parametric approach. We therefore rely on the fact that, in the standard NEG models (with no quality variation across varieties), s_{gc} and s_{bgc} are proportional to the number of varieties available and the fact that both parametric and non-parametric estimation yield very similar estimates of the counts of varieties to motivate our use of non-parametric estimates for s_{gc} and s_{bgc} .

Initially, let us assume that each household selects only one UPC out of the S UPCs available in the market. If we also assume that each UPC is purchased with a probability π that is identical for all UPCs in the market, then it follows that $\pi = 1/S$. Our task, now, is to estimate S using the number of different UPCs purchased by a sample of H households.

To do this, we make one additional assumption: stores have sufficient inventories of goods so that the purchase of a UPC by one household does not reduce the probability of another household buying the same UPC. If household purchases are independent in the cross-section, then the probability that we observe one of the H households in our sample selecting a particular UPC is equal to one minus the probability that none of the H households selects the UPC, or $1 - (1 - 1/S)^H$. The number of different UPCs that we expect to observe in the purchase records of the H households is simply the sum of these probabilities across all of the S available UPCs,

$$S(H) = S[1 - (1 - 1/S)^H] \tag{2.6}$$

It would be straightforward to obtain an estimate for the number of varieties in the market in this simple case. By equating $S(H)$ to the sample UPC count, $\tilde{S}(H)$, we can derive an estimate for the number of available UPCs, $\hat{S}$, which satisfies equation (2.6). Note that the distribution of $S(H)$ should follow the negative exponential function,

$$S(H) = S \left(1 - e^{- \ln((1-1/S)H)} \right) \tag{2.7}$$

This simple approach cannot be applied to the data for two reasons. First, households purchase more than one UPC in the course of a year. And second, some UPCs, like milk, are likely to be purchased at higher frequencies than other UPCs, like salt. Hence the probability that we observe the purchase of a UPC will vary across UPCs. We can deal with the first problem by allowing the purchase probability π to differ from $1/S$. The second problem, however, is more complicated because solving it requires us to know the purchase frequencies not only of every observed UPC but also of the UPCs that we do not observe in our sample. In order to solve this problem, we follow Mao, Colwell, and Chang (2005). We allow for different products to have different purchase probabilities, but we restrict these probabilities to be identical for groups of UPCs that we will refer to as "incidence groups." Grouping UPCs into this way enables us to estimate the purchase frequencies for all of the UPCs in an incidence group even though we may only observe one of these UPCs being purchased.

Suppose that each UPC u has a probability of π_{cu} of being selected by any one household in city c, and that there are K different incidence groups, or values that π_{cu} can take in each city c, such that $\pi_{cu} = \pi_{ck}$ for each UPC in category k. We define α_{ck} as the proportion of UPCs in city that are selected with probability π_{ck}. For example, when residents of a particular city choose to buy a particular good, one might observe one purchase probability for dry goods, another one for perishable goods, a third one for cleaning supplies, etc.

We can now write the following equation relating the shares of each incidence group, the probability a UPC in that group is selected, and the total number of UPCs in the city as

$$S \sum_{k=1}^{K} \alpha_{ck} \pi_{ck} = 1$$

The number of UPCs we expect to observe in a sample of H_c households is

$$S_c(H_c) = S_c \sum_{k=1}^{K} \alpha_{ck}(1 - (1 - \pi_{ck})^{H_c}) \tag{2.8}$$

where Mao, Colwell, and Chang (2004) note that equation (2.8) can be re-written as:

$$S_c(H_c) = S_c \sum_{k=1}^{K} \alpha_{ck}(1 - \exp(C_{ck}H_c)) \text{ where } C_{ck} = -\ln(1 - \pi_{ck})$$

It is straightforward to see that equation (2.7) is a specific case of equation (2.8) in which the selection probability is identical for all UPCs, *i.e.*, $K = 1$ and $\pi = 1/S$. Equation (2.8) is therefore referred to as the "generalized negative exponential" (GNE) model.

We can now use maximum likelihood estimation to estimate S_c as well as every α_{ck}, and π_{ck} for any choice of K, but we relegate the details of that procedure to Appendix 2.A. In order to identify the correct number of incidence groups, K, we estimate the parameters of mixture distributions for a range of values for K using the conditional likelihood function. Each distribution implies a different estimate for the total number of UPCs available in a city. We then compute the Akaike Information Criterion (AIC) for each value of K in each city. We choose between the distributions by selecting the number of incidence groups for all cities equal to the K that maximizes the sum of the AICs across all cities.[8]

---

[8]We also assume that the sampling is sufficient so that we observe some UPCs purchased in each

Once we have our estimates for the α's and π's, we can use equation (2.8) to obtain an estimate of the total number of varieties as follows:

$$\hat{S}_c = \tilde{S}_c(H_c) \left[\sum_{k=1}^{K} \hat{\alpha}_{ck}(1 - (1 - \hat{\pi}_{ck})^{H_c}) \right]^{-1} \tag{2.9}$$

where variables with circumflexes represent parameter estimates and $\tilde{S}_c(H_c)$ is equal to the sample count of distinct varieties in the city. It is useful to note that when the number of sampled households in the city, H_c approaches infinity, the fact that the α's sum to one implies that our count of the number of distinct products purchased by these households becomes our estimate of the number of varieties.

The above approach is parametric in the sense that we assume the distribution of the count of varieties depends on the underlying probabilities and shares of the various incidence groups. There is also a non-parametric approach that is often used for the same purpose based on the estimation of "accumulation curves" (see Mao, Colwell, and Chang (2005, 2004)). If we define an accumulation curve, $S(n)$, as the number of distinct UPCs purchased by a sample of n households, then the total number of different varieties available equals the asymptote of the accumulation curve.

We follow a well-developed methodology for estimating these curves. We first randomly order the households in our sample for a given city. We then count the number of unique UPCs purchased by the first household in the random ordering for a city and denote this count $S(1)$. Next, we take the data for the second household in our ordering and add it to the data for the first household in a hypothetical sample and count the number of unique UPCs in this sample, denoting it $S(2)$. We continue to add households to the hypothetical sample creating a series $(S(1), S(2), S(3), \ldots, S(H_c))$, where H_c is the total number of households in our sample for city c.

One of the problems of this approach is that each accumulation curve will be sensitive to the random order in which households are added to the curve due to both random error and sample heterogeneity. Colwell and Coddington (1994) note that the random error can be reduced by randomizing the sample order R times and generating an accumulation curve,

---

incidence group. We need this assumption because we cannot say anything about the number of goods in an incidence group that are purchased with such low a probability that no one in the sample ever buys them. In other words, we can only discuss the number of available varieties for observable classes of goods.

$S_{cr(n)}$, for each random ordering indexed by r. The random error-adjusted accumulation curve is the mean of the species accumulation curves over the different randomizations,

$$\overline{S}_c(n) = \sum_{r=1}^{R} S_{cr}(n)$$

We can adjust for random error by using mean accumulation curves over 50 randomizations, *i.e.*, R = 50.[9]

An important feature of accumulation curves is that their value in the limit when n approaches infinity is an estimate of the total number of different goods available in the city.[10] Unfortunately, theory does not tell us what the distribution will be for these accumulation curves. Therefore, we will follow Jimenez-Valverde, Jimenez Mendoza, Cano, and Munguira (2006) by estimating the parameters of various functional forms and use the AIC goodness-of-fit test to choose between a range of functional forms that pass through the origin and have a positive asymptote.

We use this accumulation curve methodology to estimate the "brand-share," s_{gc}, and the "UPC-share," s_{bgc}. Just as we build an accumulation curve corresponding to the count of the different goods represented in a sample, we can also build a curve corresponding to the national market share of the different goods represented in a sample. We estimate the asymptotes of the two different share accumulation curves for each product group in each market. The first curve is used to estimate the national market share of the brand-modules

---

[9]To determine how sensitive the adjusted accumulation curve is to sample heterogeneity, Colwell and Coddington (1994) suggest comparing the adjusted accumulation curve with the accumulation curve that we would expect to draw if all of the UPCs purchased by the households in the city sample were randomly assigned to households. We use the random placement method to calculate this expected accumulation curve and its variance as:

$$\tilde{S}(n) = S_{TOT} - \sum_{u=1}^{S_{TOT}} (1 - \frac{n}{N})^{n_u}, \text{ and } \sigma^2(n) = \sum_{u=1}^{S_{TOT}} (1 - \frac{n}{N})^{n_u} - \sum_{u=1}^{S_{TOT}} (1 - \frac{n}{N})^{2n_u}$$

where S_{TOT} is the total number of UPCs recorded in the sample, n_u is the total number of households that purchase that UPC u, and N is the total number of households in the sample. If the expected curve, $\tilde{S}(n)$, rises more steeply from the origin than the mean curve, $\overline{S}(n)$, the sample heterogeneity is greater than that which could be explained by random sampling error alone. Coleman, Mares, Willig, and Hsieh (1982) state that the hypothesis of random placement will hold if the expected curve does not deviate by more than one standard deviation from the mean curve for more than one third of the sample sizes, n, and the deviations are distributed randomly across n. We checked that this is true in our data.

[10]One might think that it would be more appropriate to use the total number of households in the city instead of the asymptote to compute the number of goods available. This would be true if every good available in a city were purchased by at least one household, but not otherwise. In practice, since the number of households in a city is quite large, it does not matter whether one focuses on the asymptote or sets n equal to the number of households.

that are available in a city within a product group, s_{gc}. The second is used to estimate the national market share of the UPCs that are available in a city within each brand-module that is available in the city, s_{bgc}.

Unfortunately, we are unable to estimate one value of s_{bgc} for each brand-module b in each product group g for each city c because household samples are sometimes too small to allow us to observe more than a few purchases of UPCs within many of the brand-modules available in a city, *i.e.*, every b in B_{gc}. Since we do not want to estimate accumulation curves for brand-modules in which there are very few purchases, we pool the data in each city and estimate a common s_{bgc} for all brand-modules within each product group in a city, which we denote $\bar{s}_{gc}$. This "average" UPC-share is an estimate for the national sales of UPCs available in a city divided by the national sales of all brand-modules available in the city, *i.e.*, the weighted average within-brand national market share of the UPCs available in a city conditional on the brand-module being available in the city. We calculate this estimate using an accumulation curve of the national sales of the UPCs available in a hypothetical sample of H households as a share of the national sales of brand-modules available in the city as a whole. For the denominator, we use national product group sales multiplied by our estimate for s_{gc}. This simplifies the variety adjustment in Proposition 5, so we can calculate the EPI for city c relative to the national sample as:

$$EPI_c = \prod_{g \in G} \left[CEPI_{gc} (s_{gc})^{\frac{1}{1-\sigma_g^a}} (\bar{s}_{gc})^{\frac{1}{1-\sigma_g^w}} \right]^{w_{gc}}.$$

By rearranging terms, we can decompose the exact price index into a common exact price index for a city $CEPI_c$ and a variety adjustment, VA_c which we define below:

$$EPI_c = CEPI_c \times VA_c = \prod_{g \in G} [CEPI_{gc}]^{w_{gc}} \times \prod_{g \in G} \left[(s_{gc})^{\frac{1}{1-\sigma_g^a}} (\bar{s}_{gc})^{\frac{1}{1-\sigma_g^w}} \right]^{w_{gc}}. \quad (2.10)$$

As we mentioned earlier, Krugman (1991) does not allow for variety and, therefore, for the variety adjustment, VA_c, to vary across cities. Instead, Krugman (1991) predicts that differences in traded goods prices, or $CEPI_c$, alone makes the aggregate price index lower in larger cities. A strict, structural test of Krugman (1991) would, therefore, focus on testing whether the common exact price index, $CEPI_c$, decreases with city size. The extension

of Krugman (1991) outlined in the previous chapter makes the additional prediction that there are more varieties available in larger cities and these variety differences further lower the aggregate price index in the larger cities. We test this prediction by focusing on whether the variety adjustment component of the exact price index, VA_c, decreases with city size. The aggregate price index agglomeration mechanism in both models is tested by examining whether the exact price index, EPI_c, is decreasing in city size.

2.4   Results

We begin our analysis by exploring how prices vary with city size. One way to benchmark our results is to compare them to the unit value results we presented in (2.1). If we regress the log of the Nielsen food *unit-value* index level in a city on the log of its population, we obtain a coefficient of 0.04 (s.e. 0.01) indicating a one log unit rise in city size is associated with a four percent rise in the unit value of groceries. Table 2.2 presents results from estimating equation (2.1) where now the key difference is that we are now gauging price differences between identical products, or UPCs, sold in different cities. The first column of the table demonstrates that if we simply regress prices on city size, the coefficient on city size falls to only one quarter the magnitude obtained in the unit value regression and is only significant at the 10 percent level. This result indicates that most of the positive relationship between prices and city size in the unit value index reflects the fact that people in larger cities buy different, higher-priced varieties of goods.

These results, however, overstate the positive association between prices and city size because the prices of identical goods are likely to vary systematically with the store in which they are purchased (presumably because some stores offer better amenities) and the buyers' responsiveness to sales. Controlling for store amenities and purchaser characteristics eliminates the statistical link between city size and prices. Most of the apparently higher prices of identical goods in larger cities is due to the fact that people in larger cities shop in nicer stores.

Most surprisingly, we find that the relationship between city size and prices turns negative (although again not significantly so) once we control for zip code income and city land prices. These results are not conclusive in part because we have not aggregated the

Table 2.2: Are prices higher in larger cities?

	$Ln(P_{uzcrh})$					
Ln(Population$_c$)	0.0107*	0.0105*	0.0045	0.0045	0.0037	-0.0015
	[0.0056]	[0.0057]	[0.0035]	[0.0034]	[0.0034]	[0.00385]
Ln(Income$_z$)	-	-	-	-	0.0091***	0.0089***
	-	-	-	-	[0.0007]	[0.0006]
Ln(Land Value$_c$)	-	-	-	-	-	0.0112***
	-	-	-	-	-	[0.0035]
HH Demographic Dummies[2]	No	Yes	No	Yes	Yes	Yes
Storename/type Dummies[3]	No	No	Yes	Yes	Yes	Yes
Observations	12.7m	12.7m	12.7m	12.7m	12.7m	12.7m
R-squared	0.928	0.928	0.935	0.935	0.935	0.935

Robust standard errors in brackets. Standard errors are clustered by city.

*** p¡0.01, ** p¡0.05, * p¡0.1

Notes:

[1] P_{uzcrh} = average price paid for UPC u in store r in zipcode z by household h living in city c.

[2] Household controls include dummies for household size, male and female head of household age, marital status, race, and hispanic.

[3] Store type dummies are used instead of storename dummies if the store name is missing or sample store sales are under $100,000.

[4] Random weight goods have been dropped from the sample.

[5] All regressions have UPC fixed effects.

data appropriately to build a price index, but they strongly suggest that the result that prices rise with city size is an artifact of nicer stores locating in larger cities.

We now turn to estimating the number of varieties available in a city based on our structural GNE approach. With 49 cities, the GNE approach involves the estimation of several hundred parameters, so we do not report all the values here. The AIC indicates that UPCs tend to fall within 10 incidence groups in terms of their purchase frequency.[11] Table 2.3 summarizes these estimates across our sample of 49 cities. We see that in all cities there are few UPCs that are purchased with very high frequency – on average, one in ten thousand UPCs are purchased with a frequency of 0.5 by a household. This would correspond to about 8 UPCs in the typical city having a purchase probability of 0.5 over

---

[11]The incidence groups do not map directly into the product groups or product modules. The estimated purchase frequency associated with each UPC within a product group or brand-module will vary with the popularity of the UPC, which may be correlated with its brand, container, size, and other characteristics. The UPCs in the high frequency incidence group tend to be the most popular varieties of products that are frequently purchased, *e.g.* 12-packs of Coke cans, which are purchased by almost a third of the households in our sample. Less popular varieties of soda tend to fall into the lower purchase frequency incidence groups. These incidence groups will also be populated with the most popular UPCs in less frequently purchased product categories, *e.g.* Fleischmann's fresh cake yeast along with the more obscure varieties of soda as well as the less popular varieties of yeast.

the course of a year by a typical household. However, we also see that the vast majority of UPCs have extremely low purchase probability. 49 percent of UPCs have a purchase probability of approximately 1 in a thousand. Thus, the product space is characterized by a few UPCs with high purchase probabilities and a vast number of UPCs that are rarely purchased.

Table 2.3: Summary Statistics for GNE Parameter Estimates

Incidence	Probability of Purchase (π_{ck})		Share of UPCs (α_{ck})	
Group (k)	Mean	Standard Deviation	Mean	Standard Deviation
1	0.496	0.096	0.00009	0.00006
2	0.332	0.103	0.00041	0.00029
3	0.226	0.085	0.00116	0.00071
4	0.152	0.066	0.003	0.00181
5	0.102	0.052	0.00757	0.0047
6	0.065	0.035	0.01953	0.01045
7	0.038	0.022	0.05083	0.01821
8	0.019	0.011	0.12115	0.01879
9	0.008	0.004	0.30803	0.03277
10	0.001	0.001	0.48823	0.03465

Another way of summarizing the estimates is to examine how the probability that a UPC is purchased varies with city size. We would expect that the probability that any consumer purchases any one UPC would go down as the range of available UPCs in the city increases. Fortunately, this is easy to examine given that our GNE structure enables us to estimate the probability that a household purchases a UPC by simply calculating

$$\Pi_c = \sum_{k=1}^{K} \hat{\alpha}_{c,k} \hat{\pi}_{c,k}.$$

In Figure 2.2, we see that the estimated average probability of purchase (which uses no population data in its estimation) decreases sharply with city size. A UPC sold in our smallest city, Des Moines, has three times the probability of being purchased by any individual household as a UPC sold in New York. The fact that households in larger cities are much less likely to buy any individual UPC is strongly suggestive of the fact that that the range of UPCs available in a city is increasing with city size.

We test this directly by using the GNE parameter estimates to calculate an estimate for total number of varieties in each city and considering how this varies with city size.

Figure 2.2
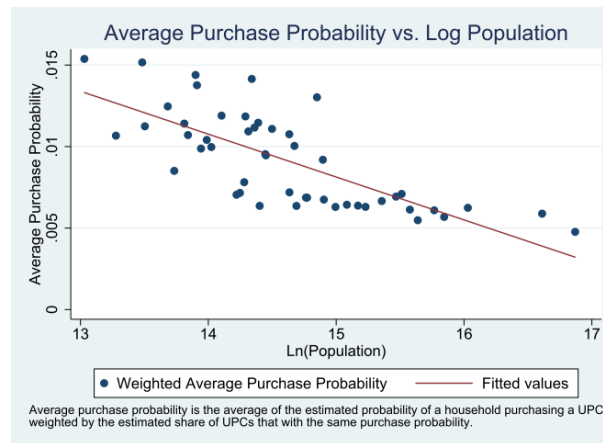
Figure 2.3 plots how the log estimated number of varieties in each city varies with city size. It is interesting that the relationship between city size and the number of varieties in a city is similar but much stronger than the one we observed in Figure 2.1. The data seem to strongly suggest a relationship between the size of a city and the number of varieties available as hypothesized by Krugman. Residents of New York have access just over 97,000 different varieties of groceries, while residents of small cities like Omaha and Des Moines have access to fewer than 32,000.

Figure 2.3
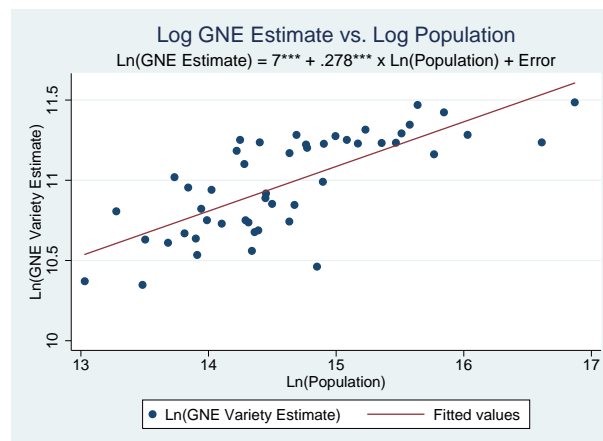
We test this relationship between city size and variety abundance formally in Table 2.4. Table 2.4 presents the results from regressing the log estimated number of varieties in a city from the GNE procedure on the log of the population in the city. The first three columns of the table present regressions of the log sample counts of varieties in each city on the log of the city's population. The next columns present regressions of the log estimate of

number of varieties based on the GNE asymptotes on city size, and the final three columns repeat these results for our Weibull estimates. As one can see, the elasticity of variety with respect to population is less using the GNE estimates because these correct for the correlation between sample size and population in the Nielsen data. What is most striking, however, is that we observe a very strong and statistically significant relationship between the size of the city and the number of estimated varieties. Our estimates indicate that a city with twice the population as another one typically has 20 percent more varieties.

Table 2.4: Do larger cities have more UPC varieties?

	Sample Count			ln(GNE Asymptote)			ln(Weibull Asymptote)		
ln(Population)	0.31***	0.34***	0.28***	0.28***	0.30***	0.26***	0.28***	0.30***	0.278***
	[0.04]	[0.07]	[0.10]	[0.04]	[0.06]	[0.08]	[0.03]	[0.05]	[0.08]
ln(Per Capita Income)	-	-0.16	-0.04	-	-0.14	-0.06	-	-0.13	-0.07
	-	[0.34]	[0.37]	-	[0.29]	[0.31]	-	[0.27]	[0.29]
Income HHI	-	-0.952	-0.29	-	-0.63	-0.18	-	-0.32	0.00
	-	[3.13]	[3.25]	-	[2.64]	[2.75]	-	[2.47]	[2.57]
Race HHI	-	0.06	0.12	-	0.07	0.11	-	0.11	0.14
	-	[0.41]	[0.42]	-	[0.35]	[0.35]	-	[0.32]	[0.33]
Birthplace HHI	-	0.01	0.03	-	-0.01	0.004	-	-0.03	-0.02
	-	[0.28]	[0.29]	-	[0.24]	[0.24]	-	[0.222]	[0.23]
ln(Land Area)	-	-	0.09	-	-	0.06	-	-	0.04
	-	-	[0.11]	-	-	[0.09]	-	-	[0.08]
Constant	6.2***	7.5**	6.3*	6.9***	8.1***	7.2**	6.9***	8.0***	7.4**
	[0.63]	[3.39]	[3.70]	[0.53]	[2.86]	[3.13]	[0.50]	[2.67]	[2.93]
Observations	49	49	49	49	49	49	49	49	49
R-squared	0.53	0.53	0.54	0.55	0.56	0.56	0.59	0.60	0.60

Standard errors in brackets; *** p¡0.01, ** p¡0.05, * p¡0.1

One concern with these results is that they might be biased because larger cities might have more diverse populations. In order to control for this we constructed a number of Herfindahl indices based on the shares of MSA population with different income, race, and country of birth. These indices will be rising in population homogeneity. In addition, we include the per capita income in each city. As one can see from Table 2.4, controlling for urban diversity does not alter the results.

Finally, we were concerned that our results might be due to a spurious correlation between city population and urban land area. If there are a constant number of unique varieties per unit area, then more populous cities might appear to have more diversity simply because they occupy more area. To make sure that this force was not driving

our results, we include the log of urban land area in our regressions. The coefficient on land area is not significant in any of the specifications, while the coefficient on population remains positive and very significant. These results indicate that controlling for land area and demographic characteristics does not qualitatively affect the strong relationship between city size and the number of available varieties. The R^2 of around 0.5 to 0.6 indicates that city size is an important determinant of variety availability. Thus, the number of tradable goods varies systematically with city size as hypothesized by our amended NEG model.

Unfortunately, our structural estimation techniques cannot be applied to the estimation of the shares because our theory of the probability of purchasing a good does not map into a theory for the probability of spending a particular share of income on a good. As we mentioned earlier, this forces us to use a non-structural technique. However, before doing so, we will first provide some intuition for the methodology and then demonstrate that non-structural methods yield almost identical results as structural methods on count data and fit share data extremely well.

We begin by plotting accumulation curves, which are a graph of the number of distinct varieties we obtain on average in a random sample of n households against the number of households. The intuition for the approach arises from the fact that if households in a city choose more disparate sets of goods than households in a second city, the accumulation curve for the first city will lie above that of the second city because n households in the first city will tend to consume more varieties than n households in the second city. One can then obtain an estimate for the total number of goods available in each city based on the asymptote of the accumulation curve for that city.

We plot these accumulation curves for the twelve cities for which we have the largest samples in Figure 2.4. These curves reveal that the four highest curves – corresponding to the cities with the greatest variety of goods purchased – are for New York, DC-Baltimore, Philadelphia, and Boston. These cities are all among the five largest cities in our sample. This raw data plot is yet another indicator that residents of larger cities consume a broader set of varieties than those in smaller cities.

We can examine this more formally by estimating the asymptotes of the accumulation curves. Since we are not sure how to model these accumulation curves, we try five

Figure 2.4
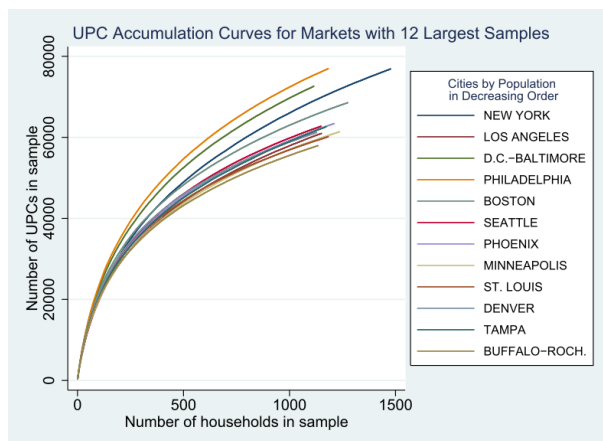
different functional forms – Clench, Chapman-Richards, Morgan-Mercer-Flodin, Negative Exponential, and Weibull. We choose among these based on the Akaike Information Criterion (AIC). The Weibull was a strong favorite with the lowest AIC score in the majority of cities for which we modeled UPC count accumulation curves, and so we decided to focus on this functional form.

Figure 2.5 plots the raw data and the estimated Weibull accumulation curve for our largest city, New York. A typical sample of 500 random households buys around 45,000 unique UPCs, and a sample of 1000 households typically purchases around 65,000 different goods. Since there is some overlap between the consumption baskets of different households in the sample, the number of unique UPCs that in the sample increases at a decreasing rate with the size of the sample. This factor produces a monotonically rising, but concave UPC accumulation curve. As one can see from the plot, the estimated Weibull distribution fits the data extremely well. The estimated asymptote is approximately 112,000 varieties, which is 35,000 more than we observe in our sample of 1500 New York households.

Figure 2.6 presents a plot of the log of the estimated Weibull asymptote against the log population in a city. As one can see, there is a clear positive relationship between the two variables: the accumulation curves imply that larger cities have more varieties than smaller ones. Despite the fact that the Weibull indicates more varieties for New York than the GNE, overall Weibull asymptotes tend to be around 20 percent lower than our GNE structural estimates. Visual inspection reveals that the relative relationship between the estimated number of varieties and city size using the GNE and plotted in Figure 2.3 is

Figure 2.5
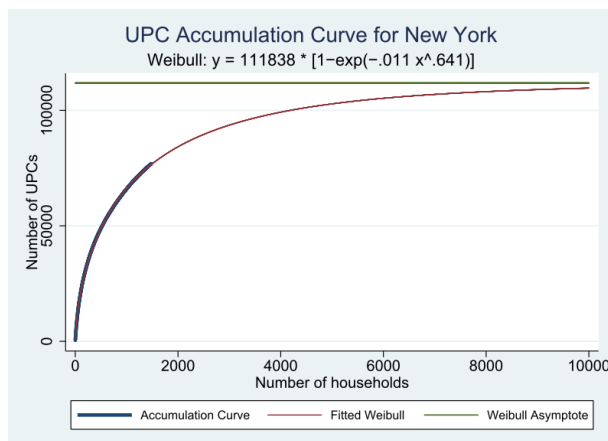
almost identical to that obtained with the Weibull. We can see this even more clearly in Figure 2.7 where we plot the Weibull asymptotes against the GNE asymptotes. The correlation between the two is 0.99 and the slope is 1.07 indicating that both methodologies yield essentially the same relationship between city size and the number of available varieties.

Figure 2.6
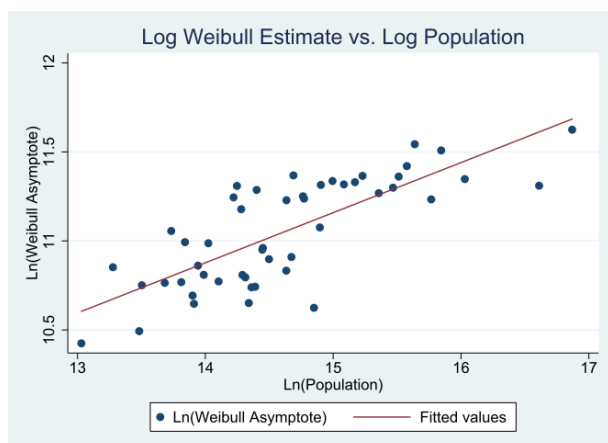
Just as we constructed accumulation curves for the number of different UPCs in cities in the previous section, we also construct these for the share of common brands in each product group, s_{gc}, and the average share of common UPCs in each brand in each product group, $\bar{s}_{gc}$. We then estimate the asymptotes of the fitted Weibull curve to each of these. Figures 2.8 and 2.9 plot these estimates for one of the largest product groups in terms of sales and the number of varieties available nationally, bread and baked goods, for two cities with large samples in the Nielsen data but very different populations, New York and

Figure 2.7
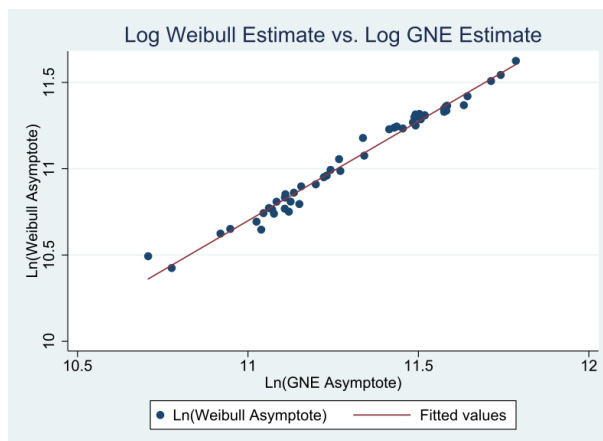
Little Rock. Figure 2.8 shows the s_{gc} accumulation curve while Figure 2.9 portrays the same curve for $\bar{s}_{gc}$. As one can see from the plots, the Weibull fits the share accumulation curves extremely well.[12] By examining the asymptote of the Weibull distribution in the first panel, we can see that New Yorkers have access to a set of bread brands available that constitute 79 percent of national expenditures on bread. By contrast, residents of Little Rock, with a population less than a tenth as large, have access to bread brands that constitute 74 percent of national expenditures. Figure 2.9 shows that the within-brand-module market share of UPCs available in a city, conditional on that brand-module being available in the city, is, on average, lower than the within-product group market share of brand-modules available in the city. Firms sell UPCs that account for 53 and 47 percent of their national sales in New York and Little Rock, respectively. This indicates that knowing that a firm sells a product in a city does not necessarily mean that all varieties of that product are available there. In summary, New Yorkers have access to 5 percent more of the bread market, in terms of brands, than residents of Little Rock and, conditional on a brand being available in their city, New Yorkers also have access to 6 percent more of the market for each brand, in terms of UPCs, than do Little Rock residents.

To demonstrate how these differences in the market shares of UPCs and brands available in the two cities translate into a discount on the exact price index in New York relative to Little Rock, we can calculate the variety adjustments for bread in each city from equation (2.10). The elasticity of substitution between UPCs within brand-modules

---

[12]The AIC favors the Weibull distribution when computing share accumulation curves as well.

Figure 2.8
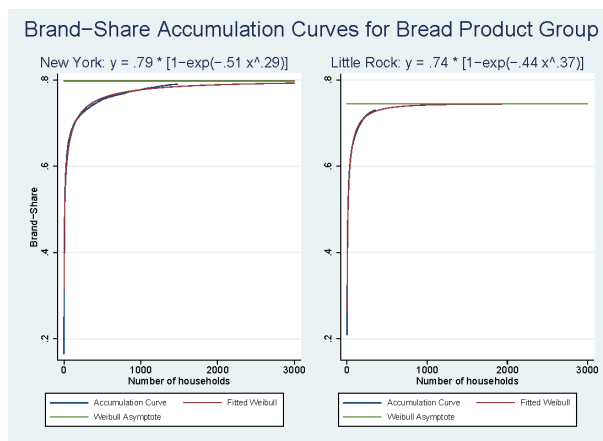
Figure 2.9
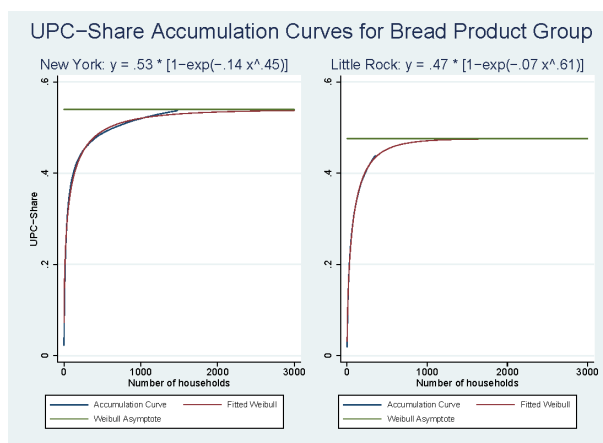
in the bread product group is 17.2, so the UPC variety adjustment for New York is $0.53^{1/(1-17.2)}$, or 1.04. The across-brand elasticity of substitution for the bread product group is 9.6, so the brand variety adjustment for New York is $0.79^{1/(1-9.6)}$, or 1.03. The fact that there are fewer bread varieties available in New York than nationally means that someone restricted to only consume those varieties available in New York would face a price index that is 7 percent higher than someone with access to all national varieties. A similar calculation shows that the variety adjustment for the bread product group in Little Rock is equal to 9 percent. The variety adjustment for New York relative to Little Rock is equal to 1.07/1.09, or 0.98, implying that if the prices of commonly available varieties are the same in the two cities, people living in New York face 2 percent lower costs for bread because they have access to more varieties.

Figures 2.10 and 2.11 plot the average asymptotes of the share accumulation curves

(estimated using a Weibull distribution) in each city against the log of the population in each city to show that the results for bread in New York and Little Rock are representative of the sample as a whole. As one can see from the figure, there is a strong positive relationship between our estimates for the national market share covered by varieties available in each city and the city's population. Nationally, consumers spend 5 percent more on brands and UPCs available in the largest cities than they do on those brands and UPCs available in the smallest cities. Once again we see that people in larger cities have access to the more and/or more popular varieties, while those in smaller cities have more limited access.

Figure 2.10
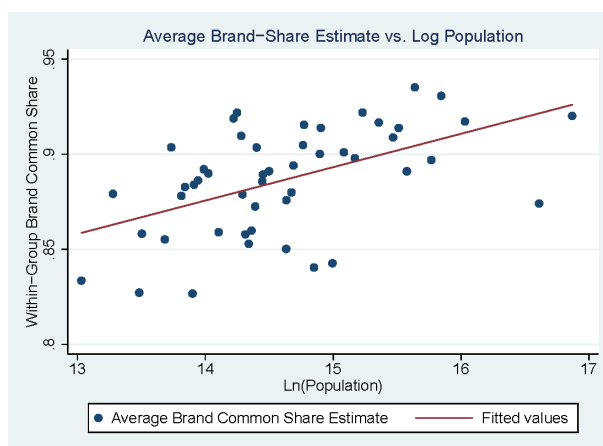
Figure 2.11
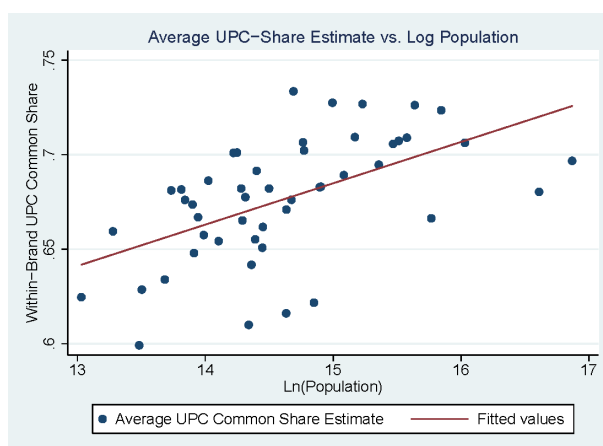
The asymptotes, in conjunction with our price data, enable us to estimate the exact price index for each product group. Table 2.5 presents our estimates for how the conventional exact price index, the variety adjustment, and the exact price index vary across

cities. The first three columns in the table use unadjusted prices to compute the index. As we saw in Table 2.2, unadjusted prices are higher in larger cities, and these results translate into a common exact price index that rises with city size. These higher prices, however, are offset by the variety adjustment to prices arising from the greater availability of varieties in larger cities, leading to an exact price index that is invariant to city size.

The unadjusted prices, however, are problematic because they do not correct for the store in which the goods are purchased or the type of household making the purchase. We therefore use adjusted prices in the subsequent columns. This yields two striking results. First, when we adjust for household characteristics in columns 4 through 6 – which control for the fact that some types of households purchase at larger discounts than others – we see that the positive relationship between city size and price ceases to be significant and the sign flips. This suggests that one important reason why we observe a positive relationship between price indices and city size is due to the fact that wealthier households congregate in large cities, and these households react less to sales.

Columns 7-9 present results in which we also control for the amenities present in the store in which the goods are purchased. These represent the fairest test of the model because we are focusing on the purchase of identical goods by similar shoppers in identical stores. The fact that there is a strong, negative relationship between the population and $CEPI_c$ provides the first empirical test of the mechanism underlying Krugman (1991). The negative and significant variety effect provides support for our augmented model. Each log unit increase in city size causes the prices of traded groceries to fall by about 2 percent and the variety adjusted index of tradable groceries to fall by 2.7 percent. These results are even more surprising considering that a factor endowments model might predict that food production would not be located in urban centers.

The remaining columns in Table 2.5 present the results of two robustness checks. Columns 10-12 show that, if we adjust prices further for the fact that people are likely to prefer shopping in nicer (*i.e.*, high income) neighborhoods, the results become even stronger. If we further purge prices of the impact of land costs and thereby look only at the tradable goods component of the price, we see in column 15 that all indices fall very sharply with city size and this result is significant at the 1 percent level.

One possible concern about these results is that people who live in large cities might

Table 2.5: Are price indices higher in larger cities?

	Unadjusted Prices			Demographic-Adjusted			Demographic-Store Adjusted			Demographic-Store-Income Adjusted			Fully-Adjusted		
	$CEPI_c$	VA_c	EPI_c	$CEPI_c$	VA_c	EPI_c	$CEPI_c$	VA_c	EPI_c	$CEPI_c$	VA_c	EPI_c	$CEPI_c$	VA_c	$EP..$
Household Effects	Included			Removed			Removed			Removed			Removed		
Store Effects	Included			Included			Removed			Removed			Removed		
Zip Income Effects	Included			Included			Included			Removed			Removed		
Land Value Effects	Included			Included			Included			Included			Removed		
Log Population	0.012**	-0.0082**	0.005	-0.014	-0.005**	-0.020*	-0.012**	-0.0051**	-0.027***	-0.026***	-0.004	-0.032***	-0.0267***	-0.005*	-0.034
	[0.005]	[0.004]	[0.007]	[0.010]	[0.003]	[0.011]	[0.008]	[0.003]	[0.009]	[0.008]	[0.003]	[0.009]	[0.008]	[0.003]	[0.00
Constant	0.825***	1.214***	1.023***	1.318***	1.143***	1.489***	1.414***	1.140***	1.589***	1.519***	1.117***	1.673***	1.521***	1.138***	1.700
	[0.076]	[0.056]	[0.11]	[0.14]	[0.037]	[0.16]	[0.13]	[0.037]	[0.14]	[0.12]	[0.038]	[0.14]	[0.13]	[0.037]	[0.1
Observations	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37
R-squared	0.138	0.12	0.012	0.054	0.117	0.09	0.137	0.109	0.189	0.23	0.059	0.263	0.226	0.103	0.28

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

Notes:

[1] The dependent variables in the above regressions are indices. These indices are calculated using unadjusted prices or prices that have been adjusted as indicated above.

[2] Random weight goods have been dropped from the sample.

[3] $EPI_c = CEPI_c VA_c$ which implies that $\log(EPI_c) = \log(CEPI_c) + \log(VA_c)$. Note that the dependent variables in the above regressions are in levels, not logs, so the coefficients on log population in the $CEPI_c$ and VA_c regressions do not add to the coefficient on log population in the EPI_c regression.

not shop in all of the neighborhoods within the city's borders. This might bias our variety adjustment downwards because we may be counting varieties in, say, suburbs that are irrelevant for people who live downtown. Alternatively, if people who live in sparsely populated neighborhoods with low variety counts often shop in densely populated neighborhoods, we might bias our variety effect upwards because measurement error leads to attenuation bias. Which bias dominates is an empirical question.

An easy solution to this problem is to make use of the fact that the Nielsen data records the household's county and the zip codes of the stores in which they purchase each UPC. We first assume that a household is located in the zip code in the county where its grocery expenditures are the highest. We then construct the population of the market where households in each zip code can choose to shop. For a set of households living in the same zip code, we sum the populations of the zip codes where we observe the set of households purchasing UPCs. This sum tells us the population of the market that is available to households located in that zip code. We average these market populations across the zip codes in each city to yield the average population of the neighborhoods where households in the city can choose to shop. We refer to this number as the relevant population.

Table 2.6 presents results in which we replace the city's population with the relevant population in the neighborhoods in which households actually shop. The results are quite similar to the results using urban population as the measure of city size. The fact that the results are so similar indicates that understanding where households shop within cities is not important for understanding the relationship between city size and prices. The key finding is that the exact price index falls with city size, and this relationship is significant whenever we adjust prices to take into account store amenities and buyer characteristics. About two-thirds of this effect appears to be due to the fact that individual goods prices are lower in larger cities, while the remainder is due to there being more, and more important, varieties available in larger cities. The economic significance of this result is fairly substantial. The demographic-store-adjusted prices in New York are 10 percent lower than those in the smallest city in our sample (Des Moines). This suggests a potentially very important role for new economic geography affecting the cost of living in cities.

Table 2.6: Are price index higher in cities? What if we adjust for varieties?

	Unadjusted Prices			Demographic-Adjusted			Demographic-Store Adjusted			Demographic-Store-Income Adjusted			Fully-Adjusted		
	$CEPI_c$	VA_c	EPI_c	$CEPI_c$	VA_c	EPI_c	$CEPI_c$	VA_c	EPI_c	$CEPI_c$	VA_c	EPI_c	$CEPI_c$	VA_c	EP
Household Effects	Included			Removed			Removed			Removed			Removed		
Store Effects	Included			Included			Removed			Removed			Removed		
Zip Income Effects	Included			Included			Included			Removed			Removed		
Land Value Effects	Included			Included			Included			Included			Removed		
Log Relevant Population	0.0079*	-0.012***	-0.0032	-0.011	-0.008***	-0.021**	-0.016***	-0.008***	-0.025***	-0.021***	-0.008***	-0.030***	-0.021***	-0.008***	-0.03
	[0.005]	[0.003]	[0.006]	[0.008]	[0.002]	[0.009]	[0.007]	[0.002]	[0.008]	[0.007]	[0.002]	[0.008]	[0.007]	[0.002]	[0.00
Constant	0.870***	1.291***	1.150***	1.306***	1.197***	1.538***	1.383***	1.194***	1.617***	1.483***	1.174***	1.700***	1.478***	1.193***	1.716
	[0.076]	[0.048]	[0.11]	[0.14]	[0.031]	[0.15]	[0.12]	[0.031]	[0.13]	[0.12]	[0.033]	[0.13]	[0.13]	[0.031]	[0.1
Observations	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37
R-squared	0.081	0.334	0.007	0.05	0.349	0.127	0.115	0.336	0.227	0.199	0.253	0.306	0.189	0.326	0.3

*** p¡0.01, ** p¡0.05, * p¡0.1

Standard errors in brackets.

Notes:

[1] The dependent variables in the above regressions are indices. These indices are calculated using unadjusted prices or prices that have been adjusted as indicated above.

[2] Random weight goods have been dropped from the sample.

[3] $EPI_c = CEPI_c VA_c$ which implies that $\log(EPI_c) = \log(CEPI_c) + \log(VA_c)$. Note that the dependent variables in the above regressions are in levels, not logs, so the coefficients on log population in the $CEPI_c$ and VA_c regressions do not add to the coefficient on log population in the EPI_c regression.

2.5 Conclusion

Krugman (1991) hypothesized that market prices drive agglomeration and the new economic geography literature which has followed rests on the same foundations. Previous empirical work using the CPI or housing costs as a cost of living measure has found that the cost-of-living is, in fact, a dis-amenity of large cities and, as such, a force of dispersion in a model of economic geography. The results presented here suggest that Krugman's result, that the price index for tradable goods is lower in larger cities, is correct. Moreover, these results support the extensions to Krugman's model in which residents of large cities consume more varieties and the two primary components of the cost of living index – housing costs and product prices – move in opposing directions as agglomeration occurs. The analysis presented here has shown that the variety-adjusted cost of grocery products decreases with city size and preliminary robustness checks using one quarter of all purchase data for a sample of 300 households in 10 cities suggest that this pattern extends throughout the traded consumer goods sector. Further work on a larger set of household purchases could confirm that the patterns identified here are not specific to grocery products.

2.A   Specifying the Likelihood Function

The starting point is to realize that the number of households purchasing product in city c, h_{cu}, follows a binomial distribution with probability:

$$P(h_{cu}) = \varphi(h_{cu}; \pi_{cu}) = \begin{pmatrix} H_c \\ h_{cu} \end{pmatrix} (\pi_{cu})^{h_{cu}} (1 - \pi_{cu})^{(H_c - h_{cu})} \tag{2.11}$$

where H_c is the total number of households in the sample for city c,

$$\begin{pmatrix} H_c \\ H_{cu} \end{pmatrix} \equiv \frac{H_c!}{h_{cu}!(H_c - h_{cu})!}$$

and once again $\pi_{cu} = \pi_{ci}$ for each UPC, u, in category k. Let $\{h_{cu}\}_{u \in U_c}$ be the observed counts of each UPC purchased by our sample of households in a city c, where U_c is the set of UPCs observed in the city c sample. Now, we can define the binomial mixture distribution as follows:

$$\Phi(h_{cu}) = \sum_{k=1}^{K} \alpha_{ck} \varphi(h_{cu}; \pi_{ck})$$

This distribution tells us the probability of observing h_{cu} purchases of any UPC u in our data, regardless of its incidence group, given the size and purchase probabilities of each of the incidence groups.

Mao, Colwell, and Chang (2005) derive a maximum likelihood methodology for estimating the α's and π's for a given K using data on the number of samples (in our case, households) in which each variety is observed. The variable n_{cj} is defined to be the number of products that are purchased by j households in the dataset for city c, *i.e.*, for which h_{cu} equals j. In other words, if 100 UPCs are purchased by no households, 50 UPCs are purchased by 1 household, and 25 UPCs are purchased by 2 households, then we would have $n_{c0} = 100$; $n_{c1} = 50$; and $n_{c2} = 25$. The joint likelihood of the total number of products available in the city, S_c, and the parameters of the mixture distribution is

$$L\left(S_c, \{\alpha_{ck}, \pi_{ck}\}_{k=1}^{K}\right) = \frac{S_c!}{\displaystyle\prod_{j=0}^{H_c} n_{cj}!} \prod_{j=0}^{H_c} \Phi(j)^{n_{cj}}$$

Note that from equation (2.8), we know that the number of available products, S_c, is a function of the number of observed products, $S_c(H_c) = \sum_{j=1}^{H_c} n_{cj}$, and the parameters of the mixture distribution, $\{\alpha_{ck}, \pi_{ck}\}_{k=1}^{K}$. Therefore, we only need to estimate the parameters of the mixture distribution to derive an estimate for the number of available products. To do so, we will maximize a conditional likelihood function. Let $\tilde{\varphi}(h_{c,u}; \pi_{c,u})$ be a zero-truncated binomial density, *i.e.*, the probability that a product is purchased by j households conditional on it being purchased by more than one household, and $\tilde{\Phi}(j) = \sum_{k=1}^{K} \alpha_{c,k} \tilde{\varphi}(j; \pi_{c,k})$ be the mixture distribution over these densities for K incidence groups. If we denote the total number of UPCs that are purchased by at least one household in the sample n_{c+}, the conditional likelihood function is

$$L\left(S_c, \{\alpha_{c,k}, \pi_{c,k}\}_{k=1}^{K}\right) = \frac{n_{c,+}!}{\prod\limits_{j=1}^{H_c} n_{c,j}!} \prod_{j=1}^{H_c} \tilde{\Phi}(j)^{n_{c,j}}$$

Chapter 3

Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities

Jessie Handbury[1]

3.1 Introduction

Previous literature measuring real wage differentials across cities and the relative cost of living across countries assumes homothetic preferences. While this assumption simplifies theory and estimation, it is contradicted by household-level studies in multiple countries showing that consumer demand varies systematically with income across a wide range of product categories.[2] This evidence has motivated economists to explore the implications of non-homotheticity for economic geography and international trade. Theory suggests that non-homotheticity might help explain the variation in returns to education across cities

---

[2]See, *e.g.*, Bils and Klenow (2001) and Broda, Liebtag, and Weinstein (2009) for the U.S., and Li (2011) for India.

(Black, Kolesnikova, and Taylor, 2009) and affect our predictions for production patterns in a closed economy setting (Wuergler, 2010) and for trade patterns in an open economy setting (see, *e.g.*, Fajgelbaum, Grossman, and Helpman (2011), Hummels and Lugovskyy (2009), and Markusen (2010)). Trade models with non-homotheticities generate new and empirically-validated predictions related to the patterns of trade (Hallak, 2006) and the prices of tradables (Simonovska, 2010). Further, Fieler (2011) has shown that a model with non-homothetic demand predicts worldwide trade patterns better than its homothetic counterpart.[3] These papers have shown that non-homotheticity is first-order in explaining trade patterns, but they do not use consumption data to identify the parameters that shape non-homothetic preferences and, therefore, cannot quantify what non-homotheticity implies for welfare in differentiated-product markets. In particular, these papers cannot compare how variation in prices and variety across markets differentially impacts the utility of consumers at different income levels. In this paper, I address this issue, making two contributions to the existing literature on non-homothetic demand.

The first contribution of this paper is methodological. Economists concerned with making aggregate statements across multiple sectors, as in trade and macroeconomics, have chiefly modeled non-homotheticities in two ways. In one approach, the perceived level of horizontal differentiation between products, which is closely related to the price elasticity of demand, varies with income such that the demand of high-income consumers is less sensitive to changes in price (see, *e.g.*, Hummels and Lugovskyy (2009) and Simonovska (2010)). In another approach, high-income consumers have a greater willingness to pay for product quality but the same sensitivity to quality-adjusted prices (see, *e.g.*, Hallak (2006) and Fajgelbaum, Grossman, and Helpman (2011)). The challenge in empirically differentiating between these forms of non-homotheticity is that quality is not clearly observed but must be imputed in some way. Empirical work that estimates non-homothetic demand for quality in multi-sector frameworks has relied on unit values as a proxy for product price and quality (Bils and Klenow, 2001; Hallak, 2006) and has therefore not been able to identify whether high-income consumers spend more on the same type of products because their demand is biased towards high quality varieties or because they are less price

---

[3]The homothetic model performs well in the context of trade between wealthy countries, but cannot reconcile the large value of trade observed between wealthy countries relative to that observed between wealthy and poor countries.

sensitive. I use both unit value and expenditure share data to estimate product quality semi-parametrically with a revealed preference approach, as in Khandelwal (2010).[4] Since I do not rely on unit values alone to measure of product quality, I can separately identify whether preferences are non-homothetic in demand for quality or in price sensitivity and test the explanatory power of a model permitting each form of non-homotheticity, either together or alone. I find that consumer purchase behavior is better explained by a model that allows for non-homotheticity in demand for quality than models that allow for non-homotheticity in both demand for quality and price sensitivity, or in price sensitivity alone.[5]

The main contribution of this paper is to show that high- and low-income consumers living in the same location reap vastly different consumption utility despite the fact that they have the same products available to them at the same prices. I measure the variations in grocery costs between U.S. cities for consumers at different income levels. Grocery costs vary with prices and product assortment across cities and this variation is different for consumers at different income levels when preferences are non-homothetic. I show that high-income households are better-off in wealthy cities because they prefer the products available in wealthy cities, even if they are sold at higher prices. Low-income households, on the other hand, get much less utility from the prices and products available in wealthier cities because the products that they prefer to consume are either more expensive or not available in these locations. This distribution of product prices and availability across cities is consistent with Waldfogel (2003)'s "preference externalities" model and the home market effects predicted in Fajgelbaum, Grossman, and Helpman (2011). The models in these papers suggest that if preferences are non-homothetic, high-income consumers will reap greater utility from living amongst other high-income consumers than low-income

---

[4]I define product quality as a product-specific taste shock that enters the log-logit utility additively. These parameters are identified by variation in product sales controlling for price, such that a product with double the sales of another product selling at the same price is estimated to have double the product quality.

[5]This result is consistent with the reduced-form results in Broda, Liebtag, and Weinstein (2009). They explain that high-income households spend more on the same type of product for two reasons. First, they buy different products that cost more per unit on average and, second, even when they purchase the same products, they pay more because they shop in more expensive stores and are less likely to purchase products at sale prices than poorer households. The authors refer to these effects as the "quality effect" and the "shopping effect," respectively. They show that the "quality effect" is ten times as large as the "shopping effect" indicating that non-homothetic demand for quality is more important than non-homothetic price sensitivity in explaining the prices consumers pay for similar products. I structurally estimate demand to confirm this prediction.

consumers do.[6] I provide the first structural estimates of the size of these consumption externalities. These estimates show the extent to which ignoring non-homotheticity biases the measurement of cross-city real wage differential estimates,[7] and suggest that assuming homotheticity may also cause large biases in cross-country welfare comparisons, such as purchasing power parity deflators. Further, they provide empirical evidence for a nascent economic geography literature that suggests that non-homotheticity is one source of skill-biased agglomeration.[8]

The model developed here is an attempt to bridge the gap between the heterogeneous agent demand models used to estimate sector-level demand in industrial organization and the representative agent multi-sector demand models of modern trade and macroeconomics. I model consumer choices in three levels. Consumers allocate expenditure between differentiated products and an outside good, among categories of differentiated products, and finally among products within each category. I model the expenditure allocation between differentiated product categories and between products within each category conditional on expenditure on a composite outside good that could include durables, education, and housing. I do not model the utility function that governs the consumer's grocery/outside good expenditure allocation decision. I assume only that the outside good is a normal good.[9] I model preferences between products within each category with

---

[6]Preference externalities, as defined by Waldfogel (2003), exist in markets for differentiated products with fixed costs of production and positive inter-city trade costs. The gathering of consumers with similar tastes, here determined by consumer income, in one location results in more production and lower prices for products suited to these tastes yielding to what Waldfogel (2003) refers to as "within-group" preference externalities. Fajgelbaum, Grossman, and Helpman (2011) show theoretically that high-income consumers gain more than low-income consumers from living in, and trading with, high-income countries when preferences are non-homothetic in the demand for quality.

[7]Other papers that have measured how relative cross-market utility varies with income include Broda and Romalis (2009) and Li (2011). Broda and Romalis (2009) calculate income-specific U.S. inflation indexes and find 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. Li (2011) shows the difference in the welfare gains from variety enjoyed by high and low consumers that is missed by ignoring non-homotheticity in demand for variety.

[8]Analysis of detailed micro-level data has shown that one of the reasons that incomes are higher in large cities is because high-skill workers co-locate in these areas (Combes, Duranton, and Gobillon, 2008). A possible explanation for this skill-biased agglomeration pattern is that high-skill workers enjoy greater productivity spillovers. However, this is inconsistent with evidence that urban wage premia decrease with skill (Adamson, Clark, and Partridge, 2004; Gyourko, Mayer, and Sinai, 2006; Lee, 2010). These papers have proposed that high-skill consumers require lower compensating wage differentials to co-locate in high-rent urban locations because they enjoy more utility from urban consumption amenities than low-skill, low-income consumers. The results presented here suggest that one source of these urban consumption amenities are preference externalities. Since consumers with similar income levels have similar tastes, they are motivated to sort geographically by income in order to maximize the local production of the products that suit their tastes.

[9]This assumption is supported in the data for the context in which I estimate demand.

a log-logit utility function. This yields continuous-discrete purchase behavior similar to that observed in the AC Nielsen HomeScan dataset.[10] I nest the within-sector log-logit utility functions in a CES superstructure that governs substitutability across products in different categories. Estimating these between-sector preferences enables me to measure aggregate utility across sectors.

Utility is non-homothetic because the amount that consumers care about the quality of differentiated products and their perceptions regarding the substitutability between these products depend on their expenditure on the outside good. The utility a consumer gets from consuming a grocery product is linear in the log quantity consumed, a product-specific parameter that is common to all consumers, and a consumer-product-specific parameter. I will refer to the product-specific parameter as product quality and the consumer-product-specific parameter as a consumer's idiosyncratic taste for the product. The weights that consumers place on the quality and idiosyncratic taste parameters vary with their expenditure on the outside good. If high-income households spend more on cars, schooling, and mortgages, for example, then they will have a greater willingness to pay for products that are ranked as high quality by all consumers; such products will yield them relatively high utility due to their idiosyncratic tastes. The willingness to pay for quality governs the elasticity of demand with respect to quality, while the willingness to pay for idiosyncratic utility is closely related to the price elasticity of demand. Both of these elasticities vary with income when the utility weights on quality and idiosyncratic tastes vary with expenditure on the outside good.

In the most general form of the model, the elasticities of demand with respect to quality and price will vary with income. These two forms of non-homotheticity have previously been modeled separately in the trade literature. Notably, in Auer (2010), utility is log-logit with consumer-specific weights on a vertical product attribute. As in the model presented here, consumers agree on the distribution of product quality, but they disagree on their willingness to pay for this vertical attribute. Auer (2010) studies the theoretical implications of this feature of the model, whereas I estimate the parameters of the model imposing that the utility weight on product quality is log-linear in expenditure on an

---

[10]The model predicts that consumers purchase a continuous quantity of a single product in each product category. This is consistent with the behavior of the typical household in the data.

outside good. This functional form assumption implies that the representative agent CES-style counterpart of the within-sector utility function is closely related to the representative agent utility function used in Hallak (2006), where the willingness to pay for quality is also log-linear in income.

The main respect in which the model here is different from Auer (2010) and Hallak (2006) is that it also permits the demand elasticity with respect to price to vary with consumer income. Consumers with higher levels of outside good expenditure place greater weight on their idiosyncratic logit utility. These logit utility draws define an "ideal type" of good for each consumer. Consumers who place greater weight on these utility draws have a greater willingness-to-pay for products close to their ideal type and, therefore, are less likely to switch to products further away from their ideal type in response to a price change. This feature is similar to the non-homotheticity in Hummels and Lugovskyy (2009). In both models, high-income consumers place a greater weight on the parameters that govern the level of horizontal differentiation between products and are, therefore, less price sensitive.[11]

Previous literature has permitted the demand elasticity with respect to product quality or price to vary with consumer income, but this is the first paper that nests both forms of non-homotheticity and estimates the parameters that govern each form of non-homotheticity independently in order to determine which model better explains consumer behavior. I estimate demand using AC Nielsen household-level scanner data containing information on the purchases of approximately 40,000 households in over 500 grocery product categories between 2003 and 2005, as well as data on the demographics of these households. I use generalized method of moments (GMM) to estimate the model under four sets of parameter restrictions that correspond to versions of the model that allow for non-homothetic demand for quality, price sensitivity, both, or neither. The parameters that govern non-homothetic demand for quality and price sensitivity are separately identified through the differences in how the expenditure shares of high- and low-income consumers vary with differences in product quality parameters and product prices, respectively. I select between the estimated versions of the model using a Bayesian Information

---

[11]This type of non-homotheticity is also generated by the translated additive-log utility function used in Simonovska (2010). More generally, this type of non-homotheticity is also generated by translated CES utility functions.

Criterion (BIC) that has been adapted for GMM by Andrews (1999). I find that the version that allows non-homotheticity in the demand for quality alone is preferred for explaining observed consumption behavior when taking into account model complexity.

I use the estimates for the selected model to calculate how a consumer's relative utility from consumption across different U.S. markets depends on his/her income. Because consumers are heterogeneous even within income groups, I summarize the utility of households with the same income level using a representative agent for each income level.[12] This analysis shows that utility from grocery consumption decreases for low-income households and increases for high-income households as they move to wealthier cities. Specifically, households earning \$15,000 per year face approximately 20 percent higher grocery costs in cities with the per capita income of San Francisco relative to cities with half that per capita income, such as New Orleans, while households earning \$100,000 per year face 20 percent lower grocery costs in the higher per capita income city.

I also measure the cross-city variation in utility predicted by the homothetic price index, calculated using the estimates of the model that does not permit non-homotheticity in either demand for quality or price sensitivity. I compare this single price index to the price indexes for different income levels and find that the grocery costs predicted by the homothetic model are highly correlated with the grocery costs I have measured for households with incomes below \$70,000, but negatively correlated with the grocery costs for households with incomes above \$100,000. This suggests that homothetic price indexes do a better job of predicting the variation in the consumption utility across cities for low-income and middle-income households than for high-income households.

This paper proceeds as follows. In Section 3.2, I introduce the dataset. In Section 3.3, I outline the model I use to estimate preferences. In Section 3.4, I outline the procedures used to estimate model parameters and demonstrate how I use the parameters to measure relative welfare across markets. In Section 3.5, I present the parameter estimates for each model and the model selection criteria implied by these estimates. Lastly, I measure how

---

[12]The utility of this representative agent is based on a nested CES counterpart to the household-specific CES-nested log-logit utility function. This representative agent utility is shown to yield the same aggregate grocery expenditure shares as the household-specific utility in an extension of Anderson, de Palma, and Thisse (1987). This proof was first extended to models that account for product quality in Verhoogen (2008). Appendix 3.C shows that the non-homothetic log-logit preferences presented have a Dixit-Stiglitz CES counterpart in that both models yield identical aggregate demand functions within groups of consumers at the same income level.

cross-city price indexes vary with income.

3.2 Data

The results in this paper are based on analysis of detailed household consumption data from the AC Nielsen HomeScan database. This data includes all food product purchases in grocery, drug, mass merchandise, and other stores for a demographically representative, but unbalanced, panel of 40,000 households in 52 cities across the United States between 2003 and 2005. The households in the sample were provided with barcode 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. AC Nielsen also surveys each household to collect information on, among other things, their income category, the number of 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 sample. AC Nielsen categorizes UPCs into 640 modules and 65 groups. The dataset includes these categorizations as well as detailed data on the brand, size (including units), container, flavor, form, formula, variety, style, organic seal, and salt content of the UPC. Of the 640 modules included in the grocery database, 46 are for random weight items and are excluded from the analysis.[13] 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.

The data includes the exact price paid by each household for each of the UPCs they purchase. Combined with the prices paid for UPCs by other households in the same market, this data is useful both for constructing market-level expenditure shares for estimation and for defining the choice set used to measure city-level grocery costs. This dataset is uniquely suited for estimating a non-homothetic utility function because it links the characteristics of the UPCs a household purchases with the demographics of the household. I discuss how I use the AC Nielsen data on product characteristics and household demographics below.

---

[13]One source of bias in the parameter estimates is unobserved correlation between the component of product quality that varies across markets or time and the prices at which these products are sold. 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.

For the structural demand estimation, I aggregate UPCs into a broader level of classification that I call a "product." One product identifier is assigned to each set of UPCs within a product module with the same brand, manufacturer, container size, salt content of a product, diet and organic categorization, and number of containers sold in a pack (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 multipack). 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." Products are defined such that this set of UPCs belong to the same product.

The utility function presented below assumes that consumers do not differentiate between UPCs in the same product. Since firms use UPCs to monitor their distribution and sales, otherwise identical products might have different UPCs because, for example, they are distributed through different channels. It is appropriate to assume that, conditional on price, consumers do not differentiate between these products. 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.

Table 3.1 shows summary statistics for the sample used for estimation. This sample has been cleaned to control for data recording errors[14] and contains 538 product modules. There are between 6 and 5,284 products in each module, and the median number of products per module is 190. The median number of UPCs per module at 289. Although there are much fewer products in each module than there are UPCs, the typical product only contains one UPC.[15]

For the purposes of this paper, the most important demographic information contained

---

[14]In order to control for data recording errors, I drop any purchase observations 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 categorization. I drop any purchase observations 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 module. I also limit the sample to products that are purchased by 20 or more households.

[15]To check the extent to which consumers differentiate between UPCs within product categories, I compare the coefficient of variation for the unit value paid for each UPC with the coefficient of variation for the unit value 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.

Table 3.1: Summary Statistics for AC Nielsen HomeScan Data Used in Estimation

Total Count	
Quarters	12
Metropolitan Statistical Areas (MSAs)	49
Modules	538
Brands	12,194
Products	181,072
Universal Product Codes (UPCs)	318,825

Count of Unique UPCs Per Category			
Category	Minimum	Median	Maximum
Module	6	289	9,464
Brand	1	2	431
Product	1	1	153

Count of Unique Products Per Category			
Category	Minimum	Median	Maximum
Module	6	190	5,284
Brand	1	2	135

Coefficient of Variation of Unit Price Paid for a UPC			
	Minimum	Median	Maximum
Within the same module	0.15	0.39	0.72
Within the same brand	0.00	0.20	1.35
Within the same product	0.00	0.15	1.34
Within the same UPC	0.00	0.14	0.98

in the dataset relates to household income. AC Nielsen classifies households into sixteen categories based on annual income. I drop the households in the three categories with annual income under \$10,000 and adjust the income of the remaining households for the number of household members. To adjust for household size, I first assign a numerical value to the income of each household. This income variable is equal to the mid-point of the bounds of a household's income category when both bounds exist and is equal to \$150,000 for those households in the "above \$100,000" income category. To adjust this income variable for the number of members in each household, I regress log household income against dummies for the ages, years of education, marital status, and race of the female and male heads-of-household, as well as fixed effects for household size. I subtract the estimated household size fixed effects from log household income and add back the one-member household fixed effect to all observations to get a projection of the log household income of each household if it were to have only one member. The distribution of the single-member household income projections are shown in Figure 3.1. The bulk of the distribution is between \$10,000 and \$80,000, which seems reasonable given that the per capita incomes of the MSAs represented in the sample range from \$21,446 in New Orleans to \$54,191 in San Francisco.

Figure 3.1: Distribution of Size-Adjusted Household Income
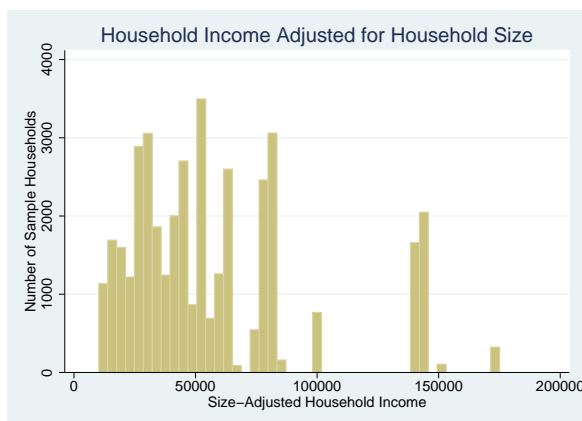
To supplement the AC Nielsen data, I also use two forms of U.S. census data. From the 2007 economic census, I use county-level food manufacturing output, the number of grocery stores, and the number of grocery store employees to create price instruments that

help identify the parameters of the model. From the 2000 U.S. census, I use per capita income and population data for each Metropolitan Statistical Area (MSA) in the sample to test whether high-income households face lower costs in wealthier and/or larger cities relative to low-income households.

Table 3.2 shows the number of households included in the sample, the population, and the per capita income for the 49 MSAs considered in the cross-city grocery cost analysis. Although per capita income is 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. This variation will help to separately 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.

3.3   Model

I will introduce the model by first outlining some notation and the general framework of consumer choices. I will then discuss the key feature of the model, non-homotheticity, which is achieved by allowing the outside good expenditure to enter the utility function of the "inside" goods, *i.e.* differentiated grocery products. Finally, I will present the function that characterizes consumer utility over grocery products and solve the consumer's utility maximization problem. I will use these solutions to generate functions for income-specific market shares that will be used to empirically identify the parameters of the model as described in Section 3.4.1 below.

3.3.1   Notation

Each consumer i earns income Y_i that he/she allocates between a set of grocery products, denoted by $\mathbf{G}$, and a set of other goods, denoted by $\mathbf{Z}$, spending W and Z in each sector, respectively. Figure 3.2 shows the consumer's expenditure allocation decision. Each oval represents the product or the set of products listed within the oval, and the arrow pointing

Table 3.2: Summary Statistics on Market Size and Income

Market ID	Market Name	Household Count	Population	Per Capita Income
1	Des Moines	145	456,022	37,650
2	Little Rock	351	583,845	33,289
3	Omaha	116	716,998	37,869
4	Syracuse	164	732,117	31,445
5	Albany	138	875,583	24,811
6	Birmingham	570	921,106	35,448
7	Richmond	194	996,512	37,082
8	Louisville	454	1,025,598	34,162
9	Grand Rapids	198	1,088,514	31,966
10	Jacksonville	130	1,100,491	35,439
11	Memphis	331	1,135,614	34,052
12	Raleigh-Durham	239	1,187,941	35,585
13	Nashville	403	1,231,311	36,056
14	Salt Lake City	248	1,333,914	33,469
15	Charlotte	1,115	1,499,293	36,580
16	Columbus	1,101	1,540,157	34,777
17	San Antonio	921	1,592,383	31,189
18	Indianapolis	245	1,607,486	36,160
19	Orlando	247	1,644,561	31,828
20	Milwaukee	170	1,689,572	37,361
21	Hartford-New Haven	193	1,725,259	41,705
22	Kansas City	220	1,776,062	35,839
23	Sacramento	1,052	1,796,857	35,318
24	New Orleans-Mobile	409	1,877,984	21,446
25	Oklahoma City-Tulsa	528	1,886,581	34,068
26	Cincinnati	317	1,979,202	35,326
27	Portland, Or	263	2,265,223	34,921
28	Buffalo-Rochester	1,135	2,268,312	33,020
29	Pittsburgh	303	2,358,695	36,159
30	Tampa	1,173	2,395,997	33,678
31	Denver	1,128	2,581,506	42,476
32	St. Louis	1,183	2,603,607	35,991
33	San Diego	139	2,813,833	40,383
34	Cleveland	362	2,945,831	35,555
35	Minneapolis	1,236	2,968,806	42,457
36	Phoenix	1,211	3,251,876	32,660
37	Seattle	1,149	3,554,760	42,356
38	Miami	1,086	3,876,380	38,342
39	Atlanta	1,038	4,112,198	35,262
40	Houston	1,070	4,669,571	40,734
41	Dallas	899	5,221,801	38,089
42	Detroit	1,001	5,456,428	37,204
43	Boston	1,275	5,819,100	47,491
44	Philadelphia	1,183	6,188,463	40,948
45	San Francisco	995	7,039,362	54,191
46	Washington, DC-Baltimore	1,115	7,608,070	46,725
47	Chicago	1,064	9,157,540	39,454
48	Los Angeles	1,151	16,373,645	37,441
49	New York	1,477	21,199,865	46,221

towards each oval represents the consumer's expenditure allocation towards that product or set of products. For example, the oval labeled "Grocery **G**" represents the set of all grocery products, and W is the expenditure that the consumer allocates to this set of products.

Figure 3.2: Consumer Choices
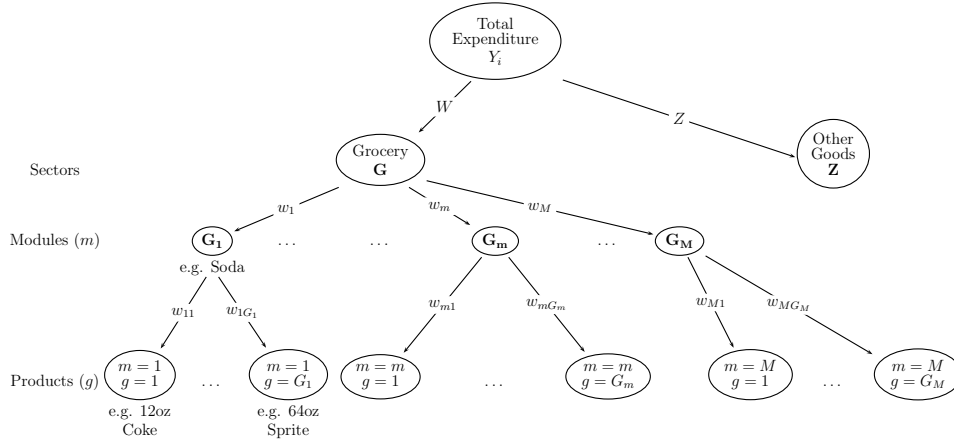
The focus of this paper will be on the choices that consumers make within the grocery sector; that is, how consumers allocate their grocery expenditure W between the varieties of grocery products available to them. The consumer's within-grocery expenditure allocation can be organized into two levels. First, consumers allocate their grocery expenditure W between distinct subsets of products called "product modules." These subsets are defined using the product module variable provided in the AC Nielsen data. One such product module is "SOFT DRINKS - CARBONATED", and a product within this module is a 12-ounce can of Coca-Cola Classic.[16] I will refer to the set of product modules as **M** and will index modules with the subscript $m = 1, ..., M$. As shown in Figure 3.2, consumers allocate some expenditure w_m to products in module m, under the constraint that their module expenditures sum to their total grocery expenditure allocation W; that is, $\sum_{m \in \mathbf{M}} w_m = W$.

The consumer's module expenditure is shared between the products in a module. I denote the set of products in a module m as $\mathbf{G_m}$. I identify products in a module by the module index, m, and a product index, $g = 1, \ldots, G_m$. The product index uniquely

---

[16]See Section 3.2 for greater detail on the product categorizations in the AC Nielsen dataset and how they are used to define products for the purpose of this analysis.

identifies products within modules, but does not across modules. A consumer chooses to spend some w_{mg} on each product g in module m. The consumer purchases $q_{mg} = w_{mg}/p_{mg}$ units of product g in module m, where units are module-specific and p_{mg} is the price of of product g in module m. I will denote the set of observed grocery prices and consumption quantities for module m as $\mathbb{P}_m = \{p_{mg}\}_{g \in \mathbf{G_m}}$ and $\mathbb{Q}_m = \{q_{mg}\}_{g \in \mathbf{G_m}}$, respectively, and the set of all grocery prices and consumption quantities as $\mathbb{P} = \{\mathbb{P}_m\}_{m \in \mathbf{M}}$ and $\mathbb{Q} = \{\mathbb{Q}_m\}_{m \in \mathbf{M}}$, respectively.

The consumer's across-module and within-module expenditure allocation decisions are linked by the fact that the consumer cannot allocate more than their total module expenditure, w_m, between products $g \in \mathbf{G_m}$; that is, $\sum_{g \in \mathbf{G_m}} w_{mg} = w_m$. If I define the consumer's within-module expenditure share as $s_{mg} = w_{mg}/w_m$, then this within-module expenditure allocation constraint is equivalent to the restriction that these shares sum to one within a module; that is, $\sum_{g \in \mathbf{G_m}} s_{mg} = 1$.

I will denote the set of all existing grocery products as $\mathbf{G}$, equal to the union of the module product sets, $\mathbf{G_m}$, over all modules $m = 1, ..., M$. The products in set $\mathbf{G}$ will be indexed by mg, where m indicates the module under which the product is characterized and g indicates the product's index within the module. The consumer's grocery expenditure allocation decision is, therefore, to allocate grocery expenditure W between products in the set $\mathbf{G}$ such that $\sum_{mg \in \mathbf{G}} w_{mg} = \sum_{m \in \mathbf{M}} \sum_{g \in \mathbf{G_m}} w_{mg} = W = Y_i - Z$, where the final equality is due to the consumer's budget constraint in their grocery/non-grocery expenditure allocation decision. A consumer allocates his/her grocery expenditure by maximizing a utility function over all grocery products $\mathbf{G}$ subject to the constraint $\sum_{mg \in \mathbf{G}} w_{mg} = \sum_{m \in \mathbf{M}} \sum_{g \in \mathbf{G_m}} w_{mg} = W = Y_i - Z$. Before explicitly stating this utility function, I will highlight the key feature that will make it non-homothetic.

3.3.2 Non-Homotheticity Through the Outside Good

For the purposes of this paper, the most important feature of the utility function presented below is that it is generically non-homothetic, *i.e.* a consumer i's expenditure allocation between grocery products depends on their total expenditure, or income, Y_i. This non-homotheticity is due to the fact that tastes for grocery products depend on the expenditure

on non-grocery products, Z. The details of this dependence, and the parameter values for which it does and does not hold, will be explained in Section 3.3.3. For now, assume that consumers get utility from consumption of the outside good and grocery products. I will denote this utility as $f(U_{iG}(\mathbb{Q}, Z), Z)$. It is a function of two terms: non-grocery expenditure, Z, and a quantity index, $U_{iG}(\mathbb{Q}, Z)$. The quantity index represents utility from grocery product consumption, $\mathbb{Q}$, which itself depends on non-grocery expenditure.

Non-grocery expenditure will be selected to maximize $f(U_{iG}(\mathbb{Q}, Z), Z)$ conditional on a budget constraint defined by consumer income and prices:

$$\max_{\mathbb{Q}, Z} f(U_{iG}(\mathbb{Q}, Z), Z) \quad \text{s.t.} \quad \sum_{m \in \mathbf{M}} \sum_{g \in \mathbf{G_m}} p_{mg} q_{mg} + Z \leq Y_i, \ q_{mg} \geq 0 \ \forall \ mg \in \mathbf{G} \qquad (3.1)$$

I do not model the consumer choice between grocery and non-grocery expenditure; however, I assume that $f(U_{iG}(\mathbb{Q}, Z), Z)$ is such that optimal non-grocery expenditure is increasing in income.

The assumption that the non-grocery composite good is normal is important for two reasons. First, it is sufficient to make preferences generically non-homothetic. Grocery utility depends on non-grocery consumption; however, this will not imply that preferences over grocery products will be non-homothetic unless optimal grocery expenditure varies with income. Second, it is necessary for an approximation used in estimation. In the analysis below, I derive estimating equations that summarize grocery consumption decisions conditional on household non-grocery expenditures, Z_i.[17] I use these equations to construct moments that are calculated using household income, Y_i, as a proxy for optimal non-grocery expenditure, Z_i^*.[18] This approximation is valid if optimal non-grocery consumption is increasing in household income.

In Section 3.B of the Appendix, I solve for an implicit restriction on $f(U_{iG}(\mathbb{Q}, Z), Z)$, prices, and model parameters under which the optimal non-grocery expenditure, Z_i^*, will be increasing in income. Although I cannot show that this restriction holds generally, I am

---

[17]Specifically, I solve for the grocery product expenditure shares that maximize a consumer's utility from grocery products, $U_{iG}(\mathbb{Q}, Z_i)$, conditional on their outside good expenditure, Z_i^*. I then estimate the parameters of the function $U_{iG}(\mathbb{Q}, Z)$ by minimizing the distance between the predicted expenditure shares of a group of households with the same optimal non-grocery expenditure, Z^*, and the expenditure shares observed in the data for groups of consumers with different levels of outside good expenditures.

[18]As a robustness check, I estimate the model using annual income less annualized observed grocery expenditures as a direct measure of non-grocery expenditure. The parameter estimates using this measure of Z are very similar.

able to show that it holds in the data. To do so, I annualize the observed grocery expenditure for each household and measure annual non-grocery expenditures as the difference between the median household income of each household's reported income category and the household's annual grocery expenditures. The elasticity of non-grocery expenditures, Z_i, with respect to household income, Y_i, is 1.05 with a standard error of 0.0003. There is also an Engel curve relationship between grocery expenditures and income. The median ratio of grocery expenditures to household income decreasing from 0.15 in the lowest income category of households to 0.05 in the highest income category. This evidence supports the assumption that non-grocery expenditure is increasing in household income, and thus justifies the use of household income as a proxy for non-grocery expenditure in estimation. In the discussion below, I will refer to households with high incomes and those with high outside good expenditures interchangeably.

3.3.3 Functional Form Assumptions

I model consumer demand for the products in $\mathbf{G}$ using a combination of constant elasticity of substitution (CES) and log-logit preferences. A consumer i's utility from grocery consumption, conditional on their outside good expenditure Z is defined as:

$$U_{iG}(\mathbb{Q}, Z) = \left\{ \sum_{m \in M} u_{im}(\mathbb{Q}_m, Z)^{\frac{\sigma(Z)-1}{\sigma(Z)}} \right\}^{\frac{\sigma(Z)}{\sigma(Z)-1}} \tag{3.2}$$

where $\sigma(Z) > 1$ is the elasticity of substitution between modules and $u_{im}(\mathbb{Q}_m, Z)$ is consumer i's utility from consumption in module m, defined as:

$$u_{im}(\mathbb{Q}_m, Z) = \sum_{g \in \mathbf{G_m}} q_{mg} \exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img}) \tag{3.3}$$

Here, q_{mg} is the consumed quantity of product g in module m, $b_{mg}(Z)$ is the taste that all consumers with outside good expenditure Z have for product g in module m, and ε_{img} is the idiosyncratic utility of consumer i from product g in module m. This idiosyncratic utility is an independent draw from a type I extreme value distribution with shape parameter 0 and scale parameter 1. $\mu_m(Z) > 0$ is the weight that consumers with outside good expenditure Z place on their idiosyncratic utility draw and thus governs the importance

that these consumers place on the horizontal differentiation between products in each module. $b_{mg}(Z)$, $\mu_m(Z)$, and $\sigma(Z)$ depend on the consumer's outside good expenditure, Z.

In what proceeds, I discuss the reasoning behind the log-logit and CES functional form assumptions made in generating the framework above. I also outline the functional form assumptions made on the $b_{mg}(Z)$, $\mu_m(Z)$, and $\sigma(Z)$ terms, explain how they yield non-homothetic demand for product quality, cross-product variety, and cross-module variety, respectively. Finally, I discuss how these functional forms relate to those used in the previous literature to model non-homothetic demand over differentiated products.

The aim of this paper is to estimate the parameters of the utility function and measure aggregate utility over many modules. The consumer's preferences across products in different modules are governed by the CES aggregator shown in equation 3.2. Since the cross-module substitution patterns are governed by a CES utility function, the model predicts that consumers will optimally spend a positive amount in each module. This prediction does not match the data, although models that yield more realistic cross-module consumption patterns would be difficult to estimate given the dimensions of the problem that this paper addresses.[19]

I allow for the substitutability of products in different modules to vary with outside good expenditure through $\sigma(Z)$. Specifically, I impose that the elasticity of substitution between modules is log-linear in outside good expenditure, $Z = \exp(z)$, such that $\sigma(Z) = 1 + \alpha^0 + \alpha^1 z$. This implies that for $\alpha^1 < 0$, consumers with high outside good expenditures, or high income, find products in different modules less substitutable for one another and will, therefore, switch between modules less than consumers with low outside good expenditures in response to changes in the relative price level across modules.

---

[19]Specifically, the typical household spends a positive amount in only 190 of 538 modules. The expenditure pattern could be replicated more closely using a less-restrictive model of multiple-discrete purchasing behavior that would allow consumers to purchase products in some, but not all, modules. For example, one could assume nested-logit preferences across all products in **G** with module-level nesting. Under this assumption, consumers would purchase multiple units of a variety of products on each purchase occasion because their purchases are intended to cover multiple consumption occasions, as in Hendel (1999) and Dube (2004). Consumer purchasing decisions could be also modeled in a static setting by assuming that consumers are endowed with a certain level of utility in each module. This method, used in Song and Chintagunta (2007) and Pinjari and Bhat (2010), implies that a consumer's maximum marginal utility from expenditure in a module must be greater than a certain cut-off value for a consumer to want to purchase any products in that module. The maximum likelihood methods these authors use to estimate these models are not computationally feasible given the size of the choice set and the number of households in the dataset.

We now restrict our focus to the log-logit sub-utility defined in equation (3.3). I refer to this as the consumer i's within-module utility function. The log-logit functional form was selected because it predicts household-level purchasing behavior that is similar to the observed pattern of household grocery purchases at the micro level. The typical household in the data purchases one or more units of exactly one product per module in each quarter.[20] The utility function provided above matches this feature of the data well. To see this, note that consumer i's module-level utility is additive in the utility he/she gains from each product g in each module m, $u_{img}(Z) = q_{mg} \exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})$. As a consequence of this additivity, these product-level utilities are perfectly substitutable with the utility from each of the other products within the same module. This implies that households will purchase positive quantities of only the product(s) that maximize their marginal utility from product expenditure, $\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})/p_{mg}$. Further, each household's product-specific idiosyncratic utility draws, ε_{img}, are drawn from a continuous distribution, so there will be a unique product that maximizes the marginal utility of expenditure for each household within each module. The module-level utility function therefore predicts that each household will purchase only one product within each module.

I parametrize the taste parameter, $b_{mg}(Z)$, which is both consumer and product specific, by assuming that a consumer's taste for product g in module m is equal to a product attribute, β_{mg}, scaled by a module-specific valuation for this attribute, $\gamma_m(Z)$; *i.e.* $b_{mg}(Z) = \beta_{mg}\gamma_m(Z)$.[21] I interpret the product attribute β_{mg} as a measure of product quality. These quality parameters are determined in the data as the average willingness to pay for product g across consumers. β_{mg} will be high for a product g in module m relative to $\beta_{m\tilde{g}}$ for another product $\tilde{g}$ in the same module m when a set of consumers

---

[20]That is, the median number of products purchased in a module by a household in a quarter is equal to one.

[21]The log-logit utility function defined in equation (3.3) 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. In Auer (2010), consumer utility is Cobb-Douglas in an outside good and a composite utility over a set of products **J** that takes the form:

$$u_i\left(\{q_g\}_{g\in\mathbf{J}}\right) = \sum_{g\in\mathbf{J}} q_g \exp(\beta_g v_i + \mu\varepsilon_{ig}) \tag{3.4}$$

Here, q_g is the consumed quantity of product g, ε_{ig} is a draw from a type I extreme value distribution with scale parameter 0 and shape parameter 1; β_g is a product attribute, v_i is consumer i's valuation of a product attribution; and $\mu > 0$ is a shape parameter that governs the level of product differentiation in the set **J**. My utility function is a generalization of Auer (2010) in that Auer restricts the weight on the idiosyncratic product utility, μ, be identical across consumers. Auer (2010) allows for a flexible interpretation of what governs the product attribute and valuation terms.

facing the same price for both products spends a higher share of their module expenditure on product g than on product $\tilde{g}$. I interpret $\gamma_m(Z)$ as the valuation for product quality, β_{mg}, for product g in module m shared by consumers spending Z on the outside good. I assume that $\gamma_m(Z)$ is log-linear in outside good expenditure, $Z = \exp(z)$, with a module specific slope, γ_m, such that:

$$b_{img} = \beta_{mg} v_{im} = \beta_{mg}(1 + \gamma_m z) \tag{3.5}$$

A consumer's valuation for product quality in module m is increasing in z for $\gamma_m > 0$, decreasing in z for $\gamma_m < 0$, and invariant to z for $\gamma_m = 0$. Under the assumption that Z is a normal good, z will be increasing in consumer income Y_i, so the weight a consumer places on product quality will also be increasing with income for $\gamma_m > 0$. Therefore, γ_m will be positive when the module expenditure share on high-quality products is higher for consumers with higher log income, y_i.

The within-module utility function defined in equation (3.3) is also non-homothetic through the weight, $\mu_m(Z)$, on the idiosyncratic utility, ε_{img}. I assume that the perceived level of horizontal differentiation between products varies with outside good expenditure $Z = \exp(z)$ according to the following equation:

$$\mu_m(z) = \frac{1}{\alpha_m^0 + \alpha_m^1 z} \tag{3.6}$$

These idiosyncratic utility weights govern the dis-utility from consuming products that are horizontally differentiated from the consumer's ideal type of product. For $\alpha_m^1 < 0$, $\mu_m(z)$ increases 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. $\mu_m(Z)$ is equal to $\frac{1}{\sigma_m(Z)-1}$, where $\sigma_m(Z)$ is the elasticity of substitution between similar quality products in module m for a consumer with outside good expenditure Z. With the functional form assumption in equation (3.6), this relationship implies that $\sigma_m(z) = 1 + \alpha_m^0 + \alpha_m^1 z$. That is, for $\alpha_m^1 < 0$, the elasticity of substitution is decreasing in consumer income. Further, consumers who spend more on

non-grocery products have a higher perception of the horizontal differentiation between products within a module and are less price sensitive.

In the expenditure equations derived below, it is easy to see that both the idiosyncratic utility weight, $\mu_m(Z)$, and the elasticity of substitution between modules, $\sigma(Z)$, will both govern the elasticity of consumer demand with respect to price. Specifically, the weight on the idiosyncratic utility draw is equal to the inverse of the within-module price elasticity of demand, and the elasticity of substitution is one minus the between-module price elasticity of demand. In an international trade model, where markets are segmented by national borders, non-homothetic models that yield income-specific price elasticities have been used to explain the fact that companies export identical products at higher prices to high-income countries than they do to low-income countries. Hummels and Lugovskyy (2009), for example, examine international prices in 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_ω^γ for $\gamma \in [0, 1]$. This weight implies an income-specific price elasticity in a similar manner to the idiosyncratic utility weights, μ_{im}, above. Income-specific price elasticities are also generated by the translated additive-log utility function used in Simonovska (2010) and translated CES utility functions more generally. The empirical tests for whether α_m^1 or α^1 are greater than zero will test the micro-foundations of this class of models. The model selection tests between a model that allows α_m^1 to differ from zero while restricting γ_m to equal zero and a model that allows γ_m to differ from zero while restricting α_m^1 to equal zero. This model selection will provide evidence that can be used to assess the relative merits of models which allow for non-homothetic price elasticities, such as Hummels and Lugovskyy (2009) and Simonovska (2010), and models that allow for non-homothetic demand for quality, such as Hallak (2006) and Fajgelbaum, Grossman, and Helpman (2011).

3.3.4   Individual Utility Maximization Problem

The utility function defined in equation (3.2) is specific to the individual through consumer income and the consumer's idiosyncratic utility draws. In this section, I solve for consumer i's optimal demand for each product g as a function of his/her outside good expenditure, Z_i, and his/her idiosyncratic utility draws. In the next section, I aggregate across the

idiosyncratic utility draws for consumers with the same income to derive expressions for income-specific market expenditure shares that will be used to estimate the parameters that govern $b_{mg}(z_i)$ and $\mu_m(z_i)$, as well as the elasticity of substitution between modules, $\sigma(z_i)$.

The consumer solves for their optimal consumption bundle by maximizing equation (3.2) subject to the following budget and non-negativity constraints:

$$\sum_{m \in \mathbf{M}} \sum_{g \in \mathbf{G_m}} p_{mg} q_{mg} \le W_i \quad \text{and} \quad q_{mg} \ge 0 \ \forall mg \in \mathbf{G}$$

I will solve this optimization problem in two steps, assuming that the consumer draws an idiosyncratic utility ε_{ig} for each product $g \in \mathbf{G}$ prior to making their purchase decision. I first solve for the consumer's optimal module bundle, $\mathbb{Q}^*_{im}(w_m)$ as a function of the consumer's expenditure in that module, w_m. I then use the solution to this problem to solve for the consumer's optimal module expenditure, $w^*_{im}(W_i)$.

The consumer's optimal bundle, $\mathbb{Q}^*_{im}(w_m)$, will maximize the consumer's within-module utility, defined in equation (3.3), subject to the above non-negativity constraints and a budget constraint limiting expenditure on all products with a module to be less than or equal to the predetermined module expenditure, w_m:

$$
\begin{aligned}
\mathbb{Q}^*_{im}(w_m) &= \arg \max_{\mathbb{Q}_m \ge 0} \sum_{g \in G_m} q_{mg} \exp(b_{mg}(Z_i) + \mu_m(Z_i) \varepsilon_{img}) \\
\text{s.t.} \quad & \sum_{m \in \mathbf{M}} \sum_{g \in \mathbf{G_m}} p_{mg} q_{mg} \le w_m, q_{mg} \ge 0
\end{aligned}
$$

Note that varieties within the same module are perfect substitutes and thus consumers will allocate all of their module expenditure, w_m, to the product(s) that maximize the marginal utility from module expenditure. Since the idiosyncratic utility draws are from a continuous distribution, there will be a unique product, $g^*_{im} \in \mathbf{G_m}$, for each module that

maximizes this marginal utility:[22]

$$g_{im}^* = \arg\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z_i) + \mu_m(Z_i)\varepsilon_{img})}{p_{mg}} \tag{3.8}$$

The optimal product choice, g_{im}^*, is independent of the consumer's module expenditure, w_m. Since all of w_m will be allocated to g_{im}^*, consumer i's optimal quantity of g_{im}^* in module m is equal to the amount that the consumer can buy at the price they observe, i.e. $q_{img_{im}^*} = w_m/p_{mg_{im}^*}$, where $p_{mg_{im}^*}$ is the price the consumer observes for g_{im}^*. For all other products g in module m, that is, for $g \neq g_{im}^*$, consumer i's optimal quantity, q_{mg}, will be equal to zero. We can summarize this definition of the consumer's optimal module bundle $\mathbb{Q}_{im}^*(w_m)$ as follows:

$$
\begin{aligned}
\mathbb{Q}_{im}^*(w_m) &= (q_{im1}^*(w_m), \ldots, q_{imG_m}^*(w_m)) \\
\text{where } q_{img}^*(w_m) &= \begin{cases} w_m/p_{mg} & \text{if } g = \arg\max_{g \in \mathbf{G_m}} \dfrac{\exp(b_{mg}(Z_i) + \mu_m(Z_i)\varepsilon_{img})}{p_{mg}} \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{3.9}
$$

The indirect utility for consumer i in module m is the value of the within-module utility, defined in equation (3.3), calculated at the optimal module bundle, defined in equation (3.9):

$$\tilde{u}_{im}(w_m) = u_{im}\left(\mathbb{Q}_{im}^*(w_m)\right) = w_m \frac{\exp(b_{mg_{im}^*}(Z_i) + \mu_m(Z_i)\varepsilon_{img_{im}^*})}{p_{mg_{im}^*}} = w_m \max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z_i) + \mu_m(Z_i)\varepsilon_{img})}{p_{mg}} \tag{3.10}$$

I can now solve for consumer i's optimal module expenditure allocations, $\mathbf{w_i^*} = (w_{i1}^*, ..., w_{iM}^*)$, by subbing the indirect module utilities, $\tilde{u}_{im}(w_m)$, into the upper-level utility function de-

---

[22]Consumers allocate all of their module expenditure to the product(s) that maximize(s) the consumer's marginal utility from expenditure, $\frac{\exp(b_{mg}(Z_i) + \mu_m(Z_i)\varepsilon_{img})}{p_{mg}}$. Since the idiosyncratic utility parameters, ε_{img}, are independent draws from a continuous distribution and the other variables that determine a consumer's marginal utility of expenditure are exogenous to the within-module product choice, there is zero probability that the value of the marginal utility of expenditure will be equal for two products. This implies that there will be a unique product that maximizes each consumer's marginal utility from expenditure in each module.

fined in equation (3.2):

$$\mathbf{w_i^*} = (w_{i1}^*, ..., w_{iM}^*) = \underset{\sum_{m \in \mathbf{M}} w_m \leq W_i}{\arg \max} \left\{ \sum_{m \in M} [\tilde{u}_{im}(w_m)]^{\rho(Z_i)} \right\}^{\frac{1}{\rho(Z_i)}} \tag{3.11}$$

Consumer i's utility from grocery consumption is a CES aggregate over his/her indirect utility $\tilde{u}_{im}(w_m)$ from module expenditures w_m for modules $m = 1, \ldots, M$. If we substitute the solution for the indirect within-module utility for each module m, defined in equation (3.10), into the target utility in the above problem we get:

$$U_i(w_1, \ldots, w_M) = \left\{ \sum_{m \in \mathbf{M}} [\tilde{u}_{im}(w_m)]^{\rho(Z_i)} \right\}^{\frac{1}{\rho(Z_i)}} = \left\{ \sum_{m \in \mathbf{M}} \left[w_m \max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z_i) + \mu_m(Z_i)\varepsilon_{img})}{p_{mg}} \right]^{\rho(Z_i)} \right\}^{\frac{1}{\rho(Z_i)}}$$

The solution to this problem is derived in Section 3.D of the Appendix:

$$\mathbf{w_i^*} = (w_{i1}^*, ..., w_{iM}^*) \text{ where } w_{im}^* = W_i \frac{\left[\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z_i) + \mu_m(Z_i)\varepsilon_{img})}{p_{mg}} \right]^{\sigma(Z_i)-1}}{P_i(Z_i)^{1-\sigma(Z_i)}} \tag{3.12}$$

$$P_i(Z_i) = \left[\sum_{m \in \mathbf{M}} \left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z_i) + \mu_m(Z_i)\varepsilon_{img})}{p_{mg}} \right)^{\sigma(Z_i)-1} \right]^{\frac{1}{1-\sigma(Z_i)}} \tag{3.13}$$

Note that consumer i's optimal module expenditure allocation depends crucially on the consumer's idiosyncratic utility draws, $\varepsilon_\mathbf{i}$.

3.3.5 Solving for Market Shares to Obtain Estimating Equations

I denote the parameters of the utility function using a vector θ defined as:

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

where $\alpha_m = \{\alpha_m^0, \alpha_m^1\}$ and $\beta_m = \{\beta_{m1}, \ldots, \beta_{mG_m}\}$. The estimation procedure outlined in section (3.4.1) below will identify the parameter values that minimize the distance between the model's predictions for the expenditure share of each product and module in the data, conditional on consumer income, and the values of these expenditure shares

that are observed in the data for market-specific groups of households with similar income levels. In Sections 3.E and 3.F of the Appendix, I integrate the optimal consumer-specific product and module expenditure functions derived in equation (3.9) and equations (3.12) and (3.13), respectively, over the consumer's unobserved idiosyncratic utility parameters $\varepsilon_{\mathbf{i}}$. Such integration yields the model's predictions for the log market share of product g relative to a base product $\bar{g}$ (both in module m) and the log market share of module m relative to a base module $\bar{m}$. Both predictions are functions of observed product prices, observed consumer income, and parameters to be estimated.

I denote the within-module expenditure share on product g in module m for a group of households with the same outside good expenditure, $Z_i = \exp(z_i)$, facing a common vector of module prices, $\mathbb{P}_m$, as $s_{mg|m}(Z_i, \mathbb{P}_m)$. As derived in the Section 3.E of the Appendix, $s_{mg|m}(Z_i, \mathbb{P}_m)$ takes the following form:

$$s_{mg|m}(Z_i, \mathbb{P}_m) = \mathbb{E}_\varepsilon[s_{img|m}] = \frac{\exp[(\alpha_m^0 + \alpha_m^1 z_i)(\beta_{mg}(1 + \gamma_m z_i) - \ln p_{mg})]}{\displaystyle\sum_{g' \in \mathbf{G_m}} \left(\exp[(\alpha_m^0 + \alpha_m^1 z_i)(\beta_{mg'}(1 + \gamma_m z_i) - \ln p_{mg'})]\right)}$$

To obtain the log market share of product g relative to product $\bar{g}_m$, I take the difference between the log expenditure share for product g and the log expenditure share for product $\bar{g}_m$ in the same module m. I take the log of this expenditure share for product g and difference from the log expenditure share for a fixed product $\bar{g}_m$ in the same module m:[23]

$$\ln(s_{mg|m}(Z_i, \mathbb{P}_m)) - \ln(s_{m\bar{g}_m|m}(Z_i, \mathbb{P}_m)) = (\alpha_m^0 + \alpha_m^1 z_i)\left[(\beta_{mg} - \beta_{m\bar{g}_m})(1 + \gamma_m z_i) - (\ln p_{mg} - \ln p_{m\bar{g}_m})\right]$$

$$(3.14)$$

Equation (3.14) 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 in module m relative to product $\bar{g}_m$ in the same module m in terms of parameters α_m, γ_m, and $(\beta_{mg} - \beta_{m\bar{g}_m})$. This equation will be used to define moments for each product $g \neq \bar{g}_m$ in each module m that will be used to estimate all of the α_m and γ_m parameters as well as each β_{mg} parameter relative to $\beta_{m\bar{g}_m}$, i.e. $\{\beta_{mg} - \beta_{m\bar{g}_m}\}_{g \in \mathbf{G_m}}$. The remaining model

---

[23]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.

parameters, α^0, α^1, and $\{\beta_{\bar{g}_m}\}_{g \in \mathbf{G_m}}$, 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 expenditure, $Z_i = \exp(z_i)$, facing a common vector of grocery prices, $\mathbb{P}$, is derived in Section 3.F of the Appendix to be equal to:

$$\mathbb{E}_\varepsilon \left[\ln s_{im} - \ln s_{i\bar{m}} \right] = -(\alpha^0 + \alpha^1 z_i) \left[\ln V_m(z_i, \mathbb{P}_m) - \ln V_{\bar{m}}(z_i, \mathbb{P}_{\bar{\mathbf{m}}}) \right] \tag{3.15}$$

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

$$V_m(z_i, \mathbb{P}_m) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\beta_{mg}(1 + \gamma_m z_i))}{p_{mg}} \right)^{-(\alpha_m^0 + \alpha_m^1 y_i)} \right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 y_i)}} \tag{3.16}$$

Together equations (3.15) and (3.16) define the expected relative module expenditure share of a set of households with income Y_i that face prices $\mathbb{P}_m$ and $\mathbb{P}_{\bar{m}}$ in terms of parameters α^0 and α^1, as well as α_m, γ_m, β_{mg} for all $g \in G_m$, and $\alpha_{\bar{m}}$, $\gamma_{\bar{m}}$, $\beta_{\bar{m}g}$ for all $g \in G_{\bar{m}}$. I use these expressions as the basis for calculating moments for each module $m \neq \bar{m}$, which, in turn, will be used to estimate α^0 and α^1, as well as $\beta_{m\bar{g}_m}$ for each $m \in \mathbf{M}$.

3.4 Empirical Strategy

I first outline how I estimate the parameters of the model using a two-step procedure based on the relative share equations derived in Section (3.3.5) above. I then discuss how I use these parameter estimates to measure the relative utility of consumers at different income levels across the geographic markets represented in the sample.

3.4.1 Estimation

I estimate the parameters of the utility function in two steps, and I split the parameter set, θ, accordingly. Equation (3.14) defines the model's prediction for market expenditure shares on each product within each module. This equation identifies the price sensitivity parameter (α_m^0, α_m^1) and quality taste-income gradient parameter (γ_m) for each module

$m \in \mathbf{M}$. It also identifies the quality of each product, β_{mg}, in each module relative to the quality of the base product in that module, $\beta_{m\bar{g}_m}$, for all products $g \in \mathbf{G}$.[24] I denote this set of parameters by θ_1:

$$\theta_1 = \left\{ \alpha_m^0, \alpha_m^1, \gamma_m, \{\beta_{mg} - \beta_{m\bar{g}_m}\}_{g \in \mathbf{G}_m} \right\}_{m \in \mathbf{M}}$$

Equation (3.15) defines the model's prediction for market expenditure shares on each module within the grocery category. Conditional on estimates for θ_1, this equation is used to identify α^0, α^1, and $\beta_{m\bar{g}_m}$ for all $m \in \mathbf{M}$ except for the base module $\bar{m}$. I denote this set of parameters by θ_2:

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

For the purposes of estimation, I proxy outside good expenditure, Z_i, with household income, Y_i. Furthermore, I split the sample of households for each quarter-MSA market, t, into income quintiles. For each market, product, and income quintile, I calculate s_{kgt}, the within-module product g expenditure share for households in each income quintile k in market t; p_{kgt}, the unit value paid for product g by households in income quintile k in market t; and y_k, the median log income of households in income quintile k in market t. The following equation, based on equation (3.14), defines the relationship between the sample relative within-module product shares and the model's prediction for these shares:

$$\ln\left(\frac{s_{kgt}}{s_{k\bar{g}_m t}}\right) = (\alpha_m^0 + \alpha_m^1 y_k)[(\beta_{mg} - \beta_{m\bar{g}_m})(1 + \gamma_m y_k) - (\ln\left(\frac{p_{kgt}}{p_{k\bar{g}_m t}}\right))] + \nu_{kg\bar{g}_m t} \quad (3.17)$$

Here, $\bar{g}_m \in \mathbf{G_m}$ is some fixed base product for each module $m \in \mathbf{M}$, y_k is the median household income of the quintile and proxies z_i in equation (3.14), and $\nu_{kg\bar{g}_m t}$ is the error in the predicted value of the relative product shares. This error includes differences between the mean prices paid and incomes of households in quintile k in market t (p_{kgt} and y_k) and the actual prices paid and incomes earned by these households. This error also includes measurement error in the collection of the raw data.

I assume that $\mathbb{E}[\mathbf{Z}_1 \cdot \nu] = 0$ for a set of instruments $\mathbf{Z}_1$. These instruments include

---

[24]I restrict β_{mg} to be equal across all products g with the same brand name.

a set of product dummies, a set of price instruments, and interaction terms between these sets of variables and quintile income, y_k. The set of product dummies includes one dummy for each product except the base product in each module $\bar{g}_m$. I cannot use prices as instruments because they might be correlated with the error term, $\nu_{kg\bar{g}_mt}$, through unobserved market-specific product tastes. For a given market, I instrument for price using four measures based on the prices observed in surrounding geographic or temporal markets. First, I use 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.[25] These instruments are similar to those used in Hausman, Leonard, and Zona (1994) and Nevo (2000). Finally, 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 instruments are similar to those used in Asker (2004) 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.

The instruments above are intended to capture temporal cost shocks that are correlated either across geography or over time. To capture spatial shocks, such as in retail costs and the level of market competition, I use county-level statistics on the number of grocery store employees, the payroll of grocery store workers, the level of food manufacturing, and the number of grocery stores per capita. 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. For each product, market, and income quintile, I represent each of the above statistics by taking the weighted average of the county-level statistic in an MSA, weighting by the purchases of a given product that consumers in a given income quintile make in each of the counties in that MSA-quarter market. This yields four more instruments. The ninth

---

[25]The region categorizations are provided in Table 3.11 in Section 3.A of the Appendix.

and final price instrument is the number of products sold in a market for each module. This is intended to capture the level of competition in each module-market.

The second step in the estimation is based on equation (3.15), again proxying for outside good expenditures with household income and replacing z_i with y_i. I define the difference between the relative module expenditure shares in the sample, where the module m expenditure share of households in income quintile k in market t is denoted by s_{kmt}, and the model's prediction for these shares as:

$$\ln\left(\frac{s_{kmt}}{s_{k\bar{m}t}}\right) = (\alpha^0 + \alpha^1 y_k)\left[\ln V_m(y_k, \mathbb{P}_{kmt}) - \ln V_{\bar{m}}(y_k, \mathbb{P}_{k\bar{m}t})\right] + u_{km\bar{m}t} \qquad (3.18)$$

where:

$$V_m(y_k, \mathbb{P}_{kmt}) = \left[\sum_{g\in\mathbf{G_m}} \left(\frac{\exp(\beta_{mg}(1+\gamma_m y_k))}{p_{kmgt}}\right)^{-(\alpha_m^0 + \alpha_m^1 y_k)}\right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 y_k)}} \qquad (3.19)$$

The parameters estimated in the first stage of estimation, θ_1, enter the equation above through the price-adjusted quality index, or inclusive value, for each module $V_m(y_k, \mathbb{P}_{kmt})$. I construct sample moments based on the above equation using data and estimates from the first stage, $\hat{\theta}_1$, to estimate the remaining parameters, θ_2.[26] These parameters include the parameters governing the elasticity of substitution between module, α^0 and α^1, and the quality parameters for the base product in each module, $\{\beta_{m\bar{g}_m}\}_{m\in\mathbf{M}}$. All but one of these quality parameters are identified in the equations above. I, therefore normalize the quality of the base product in the base module $\bar{m}$ to zero.

Note that the inclusive value is a function of the parameters in both θ_1 and θ_2. Specifically, each product quality, β_{mg}, parameter is the the sum of $(\beta_{mg} - \beta_{m\bar{g}_m})$, a component of θ_1, estimated using equation (3.17), and $\beta_{m\bar{g}_m}$, component of θ_2. We can rewrite the

---

[26]The point estimates for θ_1, which I will denote $\hat{\theta}_1$, are used to generate right-hand side variables used in the second stage of estimation. Since the first step of estimation yields consistent point estimates for θ_1, the second step of estimation will also yield consistent point estimates for the θ_2 parameters. However, the covariance matrix for $\hat{\theta}_2$ will need to be adjusted to account for the error in the calculation of $\hat{\theta}_1$ in order to get consistent standard errors for $\hat{\theta}_2$.

inclusive value function so that it is log linear in the $\beta_{m\bar{g}_m}$ parameters to be estimated:

$$\ln V_m(y_k, \mathbb{P}_{kmt}) = \ln \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\beta_{mg}(1 + \gamma_m y_k))}{p_{kmgt}} \right)^{-(\alpha_m^0 + \alpha_m^1 y_k)} \right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 y_k)}}$$

$$= \ln \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp((\beta_{mg} - \beta_{m\bar{g}_m})(1 + \gamma_m y_k))}{p_{kmgt}} \right)^{-(\alpha_m^0 + \alpha_m^1 y_k)} \right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 y_k)}} + \beta_{m\bar{g}_m}(1 + \gamma_m y_k)$$

Under the normalization that $\beta_{\bar{m}\bar{g}_{\bar{m}}} = 0$, and using the decomposition of the inclusive value function above, we can rewrite equation (3.18) as:

$$\ln \left(\frac{s_{kmt}}{s_{k\bar{m}t}} \right) = (\alpha^0 + \alpha^1 y_k) \left[\Delta V_{1m\bar{m}}(y_k, \mathbb{P}_{kmt}, \mathbb{P}_{k\bar{m}t}, \theta_1) + \beta_{m\bar{g}_m}(1 + \gamma_m y_k) \right] + u_{km\bar{m}t},$$

where $\Delta V_{1m\bar{m}}(y_k, \mathbb{P}_{kmt}, \mathbb{P}_{k\bar{m}t}, \theta_1) = \ln V_{1m}(y_k, \mathbb{P}_{kmt}, \theta_1) - \ln V_{1\bar{m}}(y_k, \mathbb{P}_{k\bar{m}t}, \theta_1)$ and

$$V_{1m}(y_k, \mathbb{P}_{kmt}, \theta_1) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp((\beta_{mg} - \beta_{m\bar{g}_m})(1 + \gamma_m y_k))}{p_{kmgt}} \right)^{-(\alpha_m^0 + \alpha_m^1 y_k)} \right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 y_k)}}$$

The $u_{km\bar{m}t}$ errors are equal to the difference between the observe and predicted values of relative module shares in each market when the predicted values are calculated using the true values for the θ_1 parameters. In practice, the predicted values of the relative module shares in each market will be calculated using first-stage estimates for θ_1. There will be additional errors in the sample predicted values due to the fact that $\Delta V_{1m\bar{m}}(y_k, \mathbb{P}_{kmt}, \mathbb{P}_{k\bar{m}t}, \theta_1) \neq \Delta V_{1m\bar{m}}(y_k, \mathbb{P}_{kmt}, \mathbb{P}_{k\bar{m}t}, \hat{\theta}_1)$ and $\gamma_m \neq \hat{\gamma}_m$. I will denote these errors by $\nu_{km\bar{m}t}$. Taking these additional errors into account, the estimating equation defined in equations (3.18) and (3.19) becomes:

$$\ln \left(\frac{s_{kmt}}{s_{k\bar{m}t}} \right) = (\alpha^0 + \alpha^1 y_k) \left[\Delta V_{1m\bar{m}}(y_k, \mathbb{P}_{kmt}, \mathbb{P}_{k\bar{m}t}, \hat{\theta}_1) + \beta_{m\bar{g}_m}(1 + \hat{\gamma}_m y_k) \right] + \nu_{km\bar{m}t} + u_{km\bar{m}t}$$

I assume that $\mathbb{E}[\mathbf{Z}_2 \cdot (\nu + u)] = 0$ for a set of instruments $\mathbf{Z}_2$. $\mathbf{Z}_2$ includes a set of dummies for all modules except the base module, and estimates for $\Delta V_{1m\bar{m}}(y_k, \mathbb{P}_{kmt}, \hat{\theta}_1)$, and all of these instruments interacted with income, y_k.

This estimation procedure yields consistent estimates for θ_2, but the variance-covariance matrix of these parameters will be biased due to the presence of the first-step estimates

for θ_1 in the ν 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 θ_2 estimates.[27]

3.4.2   Measuring Relative Utility Across Markets

A main goal of this paper is to use the model estimates to measure how variation in the prices and product availability across U.S. cities differentially impacts the utility of consumers at different income levels. I assume that all households with the same income level spend the same amount on groceries in each market t such that $W_{it} = W_k$ for all households i with income y_k.[28] The utility that households receive from groceries in a city will still vary between households that earn the same income because their preferences depend on their idiosyncratic utility draws, the ε_{img}'s. These draws imply that no two consumers with the same income levels, or even the same outside good expenditures, have the same indirect utility from the same price vector and product set. The most direct way to summarize the indirect utilities of consumers with the same income level would be to take the expectation of the indirect utility over the idiosyncratic draws. The log-logit module-level utility function is linear in these idiosyncratic draws, making it possible to derive an analytic function for the expected module-level utility for a consumer conditional on his/her income. These module-level utilities, however, are nested within a non-linear CES aggregator, making it difficult to derive an analytic function for the expected income-specific utility from consumption in many modules. I will deal with this problem in two ways.

---

[27]This adjusted variance-covariance matrix is currently being calculated. These adjustments are required for consistent standard errors for estimates of the quality of the base product in each module, $\beta_{m\bar{g}_m}$, or the elasticity of substitution between modules, σ, but not for estimates of the remaining model parameters. The parameter estimates summarized in Section 3.5.1 are for parameters calculated in the first stage of estimation, so their statistical significance is determined using the consistent standard errors generated in the first stage of estimation. The parameters estimated in the second stage are used to calculate the cross-city price indexes, as described in Section 3.4.2 below, and their standard errors are be required to obtain consistent estimates of the measurement error in these price indexes. This measurement error can be ignored in the analysis that follows in Section 3.5.3, since the price indexes are dependent variables in these regressions.

[28]Theoretically, this assumption could be violated since consumers at each income level may choose to change their grocery expenditure (W) and outside good expenditure (Z) as they move from one city to another and thus face a different grocery price index. In the data, however, the share of income that households spend on grocery products does not vary greatly across markets (see Figure 3.11 in Appendix 3.A).

The first method is to numerically integrate over the idiosyncratic utility draws. To do this, I take N random ε_{img} draws for each product g in each module m in the sample.[29] I randomly assign one draw for each product to each of N households. I split these households into 10 groups and assign each group an income equal to the median income in each income decile of the sample households. I then calculate the indirect utility of a household from the set of products and prices they face in a market t by solving for their expenditure allocations.

To calculate a household's indirect utility, I must first determine their optimal within-module product choice by finding the product that maximizes the consumer's marginal utility from expenditure in a module. Recall that the solution to this problem is:

$$g_{imt}^* = \arg\max_{g \in \mathbf{G_{mt}}} \frac{\exp(b_{mg}(Z_i) + \mu_m(Z_i)\varepsilon_{img})}{p_{mgt}},$$

where the set of available products $\mathbf{G_{mt}}$ and product prices p_{mgt} are now specific to the market for which I am estimating a household's utility and $b_{mg}(Z_i) = \hat{\beta}_{mg}(1 + \hat{\gamma}_m y_i)$ and $\mu_m(Z_i) = \hat{\alpha}_m^0 + \hat{\alpha}_m^1 y_i$ are calculated using model estimates for β_{mg}, γ_m, α_m^0, and α_m^1, and the assumed household income y_i. The idiosyncratic utility draws, or ε_{img}'s, take the values simulated for household i. I then calculate each household's module expenditures using equations 3.12 and 3.13:

$$\mathbf{w_{it}^*} = (w_{i1t}^*, ..., w_{iMt}^*) \text{ where } w_{imt}^* = W_{it} \frac{\left[\frac{\exp(b_{mg_{imt}^*}(Z_{it}) + \mu_m(Z_{it})\varepsilon_{img_{imt}^*})}{p_{mg}}\right]^{\sigma(Z_{it})-1}}{P_i^{1-\sigma(Z_{it})}}$$

$$P_i = \left[\sum_{m \in \mathbf{M}} \left(\left(\frac{\exp(b_{mg_{imt}^*}(Z_{it}) + \mu_m(Z_{it})\varepsilon_{img_{imt}^*})}{p_{mg_{imt}^* t}}\right)^{\sigma(Z_{it})-1}\right)\right]^{\frac{1}{1-\sigma(Z_{it})}}$$

Here, $\sigma(Z_i) = 1 + \hat{\alpha}^0 + \hat{\alpha}^1 y_i$ is calculated using model estimates for α^0 and α^1 and the assumed household income y_i. Denote the utility of a household i with income y_i from grocery consumption in market t as $V_{iG}(\mathbb{P}_t, y_i)$. This utility is calculated by plugging the

---

[29]In the analysis below, I use N=5000.

optimal product choices and expenditures into the direct utility function, *i.e.*

$$V_{iG}(\mathbb{P}_t, y_i) = W_{it} \left\{ \sum_{m \in \mathbf{M}} \left[\sum_{g \in G_m} \frac{s^*_{imt}}{p_{img^*_{imt}t}} \exp(b_{mg^*_{imt}}(Z_{it}) + \mu_m(Z_{it})\varepsilon_{img^*_{imt}}) \right]^{\rho(Z_{it})} \right\}^{\frac{1}{\rho(Z_{it})}},$$

where $s^*_{imt} = w^*_{imt}/W_{it}$ is the share of grocery expenditure that household i optimally spends in module m in market t.

Recall that I have assumed that grocery expenditures, W_{it}, do not vary across cities for households with the same income Y_k, such that $W_{it} = W_k$ and $Z_{it} = Z_k$. I normalize W_k to equal one and denote the per dollar utility of a household i earning $Y_k = \exp(y_k)$ in market t with $\tilde{V}_{iG}(\mathbb{P}_t, y_k)$:

$$\tilde{V}_{iG}(\mathbb{P}_t, y_k) = \left\{ \sum_{m \in \mathbf{M}} \left[\sum_{g \in G_m} \frac{s^*_{imt}}{p_{img^*_{imt}t}} \exp(b_{mg^*_{imt}}(Z_k) + \mu_m(Z_k)\varepsilon_{img^*_{imt}}) \right]^{\rho(Z_k)} \right\}^{\frac{1}{\rho(Z_k)}},$$

I calculate the per dollar indirect utility in each market t for the set of N simulated households. I take the mean of the per dollar indirect utilities of the $N/10$ households with common income Y_k to approximate the expected per dollar utility of a household earning income Y_k in market t. I denote this expected utility using $\tilde{V}_G(\mathbb{P}_t, y_k)$:

$$\tilde{V}_G(\mathbb{P}_t, y_k) = \frac{1}{N/10} \sum_{i=1}^{N/10} \left\{ \sum_{m \in \mathbf{M}} \left[\sum_{g \in G_m} \frac{s^*_{imt}}{p_{img^*_{imt}t}} \exp(b_{mg^*_{imt}}(Z_k) + \mu_m(Z_k)\varepsilon_{img^*_{imt}}) \right]^{\rho(Z_k)} \right\}^{\frac{1}{\rho(Z_k)}}$$

The numeric integration method above is computationally intensive. Therefore, I reserve it for a robustness check and apply it only to the estimation of relative consumer utility across one market-pair. For the majority of the analysis, I measure the 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 for households with income y has utility

$$U_i = \left\{ \sum_{m \in M} \left[\sum_{g \in G_m} [q_{mg} \exp(b_{mg}(Z_i))]^{\rho_m(Z_i)} \right]^{\frac{\rho(Z_i)}{\rho_m(Z_i)}} \right\}^{\frac{1}{\rho(Z_i)}},$$

where $b_{mg}(Z_i) = \beta_{mg}(1 + \gamma_m y_i)$, $\rho_m(Z_i) = \frac{1-\sigma_m(Z_i)}{\sigma_m(Z_i)}$ for $\sigma_m(Z_i) = 1 + \alpha_m^0 + \alpha_m^1 y_i$, and $\rho(Z_i) = \frac{1-\sigma(Z_i)}{\sigma(Z_i)}$ for $\sigma(Z_i) = 1 + \alpha^0 + \alpha^1 y_i$. In Section 3.C of the Appendix, 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 for that I estimate.

Since the utility of the representative agent for households with income y_i is CES, their indirect utility is the product of their grocery expenditure and a market-specific price index that summarizes the set of prices for available products in the market. I denote this indirect utility as $V_G^{CES}(\mathbb{P}_t, y_i)$, and it is expressed as

$$V_G^{CES}(\mathbb{P}_t, y_i) = W_k P(\mathbb{P}_t, y_i),$$

where

$$P(\mathbb{P}_t, y_i) = \left[\sum_{m \in \mathbf{M}} \left(\left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(b_{mg}(Z_i))}{p_{mgt}} \right)^{(1-\sigma_m(Z_i))} \right]^{\frac{1-\sigma(Z_i)}{1-\sigma_m(Z_i)}} \right) \right]^{\frac{1}{1-\sigma(Z_i)}}$$

By assuming an exogenous W_k for the representative agent of each income category k, I have again assumed that the representative agent's grocery expenditure does not vary across markets. Under this assumption, we can measure the change in a consumer's grocery utility from one market to another as the ratio of the market- and income-specific price indexes, $\pi_{t,t'}(y_i) = P(\mathbb{P}_t, y_i)/P(\mathbb{P}_t, y_i)$. The magnitude of the relative price index in market t relative to market t' above (or below) zero indicates how much lower (or higher) the representative consumer's grocery utility is in market t relative to market t'.

3.5 Results

3.5.1 Parameter Estimates

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

Table 3.3 summarizes the estimates for the module-level parameters in each of these

four models across over 500 modules. Table 3.4 summarizes parameter estimates that
are statistically significant at the 95 percent level.[30] Columns [1] through [3] of each
table summarize the parameter estimates for the un-restricted version of the model. The
corresponding utility function that governs consumer i's preferences between products
$g \in \mathbf{G_m}$ within each module m is obtained by subbing the parametrizations for $b_{mg}(Z_i)$
and $\mu_m(Z_i)$, provided in equations 3.5 and 3.6, respectively, into equation 3.3:

$$u_{im}(\mathbb{Q}_m) = \sum_{g \in \mathbf{G_m}} q_{mg} \exp\left(\beta_{mg}(1 + \gamma_m z_i) + \frac{\varepsilon_{img}}{\alpha_m^0 + \alpha_m^1 z_i}\right), \qquad (3.20)$$

where β_{mg} characterizes the quality of product g relative to other products and ε_{img} is
consumer i's idiosyncratic utility from product g. In this model, the weights that con-
sumers place on quality and their idiosyncratic utility when determining their product
ranking are functions of their non-grocery expenditure, $Z_i = \exp(z_i)$, which is proxied in
estimation by consumer income, $Y_i = \exp(y_i)$. The estimate for the elasticity of substi-
tution between products is a function of the estimated weight placed on the idiosyncratic
utility, $\hat{\sigma}_m(y_i) = 1 + \hat{\alpha}_m^0 + \hat{\alpha}_m^1 y_i$. Therefore, when the estimated value for this weight varies
with income, or $\alpha_m^1 \neq 0$, the elasticity of substitution will also vary with income. The first
column of Table 3.3 reports the elasticity of substitution of a consumer with the mean
log income level in the sample for each module, or $\hat{\bar{\sigma}}_m = 1 + \hat{\alpha}_m^1 + \hat{\alpha}_m^1 \bar{y}_m$. 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 param-
eter estimates are well-identified in most modules, with over 380 out of 504 significant at
the 95 percent level in all four models.

   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 slightly
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 slightly

---

[30]Statistical significance implies that the lower bound of the 95 percent confidence interval for the σ_m
estimates is greater than one or that the lower (upper) bound of the 95 percent confidence intervals for
α_m^1 and γ_m estimates are greater (less than) zero.

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. All of these estimates are, however, much lower in magnitude than the own-price elasticities implied by the elasticity of substitution estimates in Broda and Weinstein (2008).[31]

Table 3.3: Summary Statistics for Parameter Estimates

Model:	NH in Quality and Price			NH in Quality		NH in Price		Homothetic
Restrictions:	None			$\alpha_m^1 = 0$		$\gamma_m = 0$		$\gamma_m = 0$ & $\alpha_m^1 = 0$
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Parameter:	$\bar{\sigma}_m$	γ_m	α_m^1	$\bar{\sigma}_m$	γ_m	$\bar{\sigma}_m$	α_m^1	$\bar{\sigma}_m$
Count	504	504	504	504	504	504	504	504
p25	1.58	-0.04	-0.19	1.62	-0.08	1.66	-0.03	1.64
p50	2.09	0.15	-0.05	2.13	0.09	2.14	0.04	2.11
p75	2.59	0.30	0.05	2.60	0.19	2.62	0.15	2.61

Table 3.4: Summary Statistics for Statistically Significant Parameter Estimates

Model:	NH in Quality and Price			NH in Quality		NH in Price		Homothetic
Restrictions:	None			$\alpha_m^1 = 0$		$\gamma_m = 0$		$\gamma_m = 0$ & $\alpha_m^1 = 0$
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Parameter:	$\bar{\sigma}_m$	γ_m	α_m^1	$\bar{\sigma}_m$	γ_m	$\bar{\sigma}_m$	α_m^1	$\bar{\sigma}_m$
Count	383	254	171	380	301	393	263	391
p25	1.87	0.15	-0.31	1.93	0.05	1.88	-0.05	1.88
p50	2.27	0.23	-0.19	2.28	0.15	2.26	0.10	2.26
p75	2.72	0.35	-0.10	2.73	0.23	2.73	0.19	2.69

Columns [2] and [5] of Table 3.3 summarize the distribution of the estimated values for γ_m across all modules. All four models assume that all consumers agree on the relative quality of products, as described by the distribution of the β_{mg} parameters for products $g \in \mathbf{G_m}$ within a module m. For positive values of γ_m, however, the utility weight that consumers place on this component of utility, relative to their idiosyncratic utility draw for each product or the quantity consumed, is increasing in their outside good expenditure z_i.

---

[31]Broda and Weinstein (2008) use the Feenstra (1994) methodology to identify the elasticity of substitution between products. This method uses the assumption that all products in the same module share the same price elasticity of supply to help identify the price elasticity of demand. 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 rise 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 β_{mg} parameters.

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 β_{mg} estimates, that is, the products for which all consumers have a higher willingness to pay. In over half of the modules represented in the data, the willingness to pay for quality is increasing in income.

Columns [2] and [5] of Table 3.4 show that over half of the estimates for γ_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 γ_m estimates are positive in the model that allows for non-homotheticity in demand for quality alone. Figure 3.4 shows that almost all of the statistically significant γ_m estimates are positive in the model that allows for non-homothetic demand for both quality and price sensitivity. These results indicate that non-homotheticity is important in many sectors. Since the demand for quality is increasing with income in most grocery sectors, richer households appreciate product quality more than poorer households.

Figure 3.3: Distribution of γ_m Parameter Estimates Across Modules
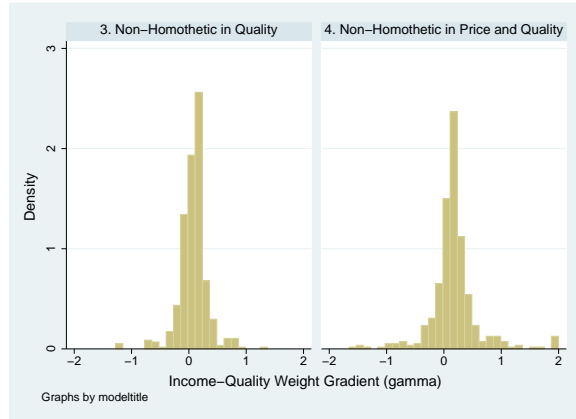
Columns [3] and [7] of Table 3.3 summarize the distribution of the estimated values for α_m^1 in each module. In equation 3.20, one can see that the weight that consumers place on their idiosyncratic utility is increasing in their outside good expenditure, z_i, for negative values of α_m^1. Suppose that two consumers draw very high values for ε_{img^*}, such

Figure 3.4: Distribution of Statistically Significant γ_m Parameter Estimates Across Modules
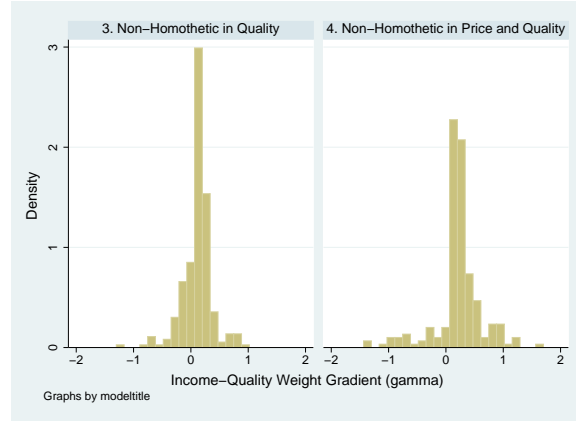
that both consumers select to consumer product g^* at the current market price. If the price of product g^* increases, then the consumer who places a higher weight on his/her idiosyncratic utility draw will be less likely to switch to another product for which he/she drew a lower value for ε_{img} relative to the product's quality, β_{mg}. For $\alpha_m^1 < 0$, 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 3.3 shows that, for the majority of modules, high-income consumers are less price sensitive, or $\hat{\alpha}_m^1 < 0$, when you control for the fact that they also have a greater willingness to pay for quality. While Column [3] of Table 3.4 indicates that 171, or fewer than half, of the 504 α_m^1 estimates are statistically significant, the right hand panel of Figure 3.6 shows that the vast majority of these statistically significant α_m^1 estimates are less than zero. While the evidence that price sensitivity varies with income is less prevalent across modules, the price sensitivity 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 Tables 3.3 and 3.4 show that the majority of the α_m^1 estimates, and even the majority of those that are statistically significant, are positive when γ_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: $\ln s_{kgt} - \ln s_{k\bar{g}_m t} = (\alpha_m^0 + \alpha_m^1 y_k)[(\beta_{mg} - \beta_{m\bar{g}_m}) - (\ln p_{kgt} - \ln p_{k\bar{g}_m t})] + \nu_{kg\bar{g}_m t}$. 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 $Corr(\beta_{kmg} - \beta_{km\bar{g}_m}, \ln p_{kgt} - \ln p_{k\bar{g}_mt}) \neq 0$, then the assumption that $\mathbb{E}[\mathbf{Z}_1 \cdot \nu] = 0$ will be violated. The fact that the α_m^1 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, $(\beta_{mg} - \beta_{m\bar{g}_m})\gamma_m y_k$, 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 α_m^1 estimates in the model that does not control for correlations in income-product specific tastes as evidence that high-income consumers are more price sensitive than low-income consumers. Instead, the positive α_m^1 estimates highlight the difficulty in identifying the non-homotheticity related to price sensitivity in isolation from the non-homotheticity related to product quality.

Figure 3.5: Distribution of α_m^1 Parameter Estimates Across Modules
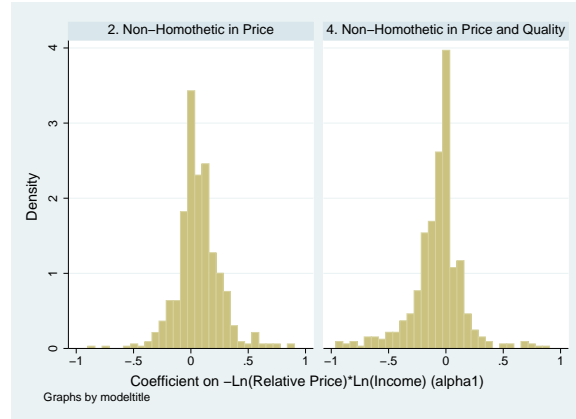
Despite the results from this one model, the parameter estimates generally show convincing evidence of non-homothetic demand. Specifically, high-income consumers have a greater willingness to pay for quality than low-income consumers and, when controlling for this non-homotheticity in the demand for quality, the results show that high-income consumers are also less price sensitive.

Figure 3.6: Distribution of Statistically Significant α_m^1 Parameter Estimates Across Modules
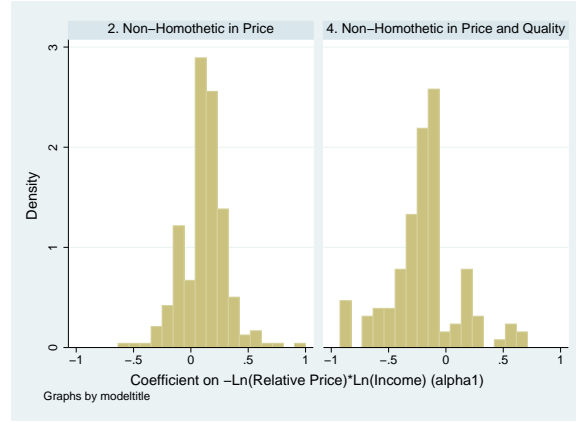
3.5.2 Model Selection

We now turn to answering the first novel question that this paper addresses: which model does the best job of explaining consumer behavior, one that allows for high-income consumers to have a greater demand for quality than low-income consumers or one that allows for high-income consumers to be less price sensitive than low-income consumers? I address this question using the GMM-BIC model selection criterion that judges models using a trade-off between model complexity, measured using the number of parameters relative to the number of moments used in the estimation of those parameters, and model fit.[32] 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. I estimate the parameters that govern the within-module product choice for each module m, denoted θ_m^1, 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 θ_m^1 are estimated. These include α_m^0, α_m^1, γ_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

---

[32]This method was developed in Andrews (1999) as a moment selection criterion and is shown to be consistent for model selection in Andrews and Lu (2001).

the demand for quality and price sensitivity. Therefore, I estimate all four models using the optimal weighting matrix for the full model, which I denote by W^*. 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. I denote this set of instruments by L^*.

I select between models in each module by minimizing the following GMM-BIC criterion:

$$\text{GMM-BIC}_{Mm}(\hat{\theta}^1_{Mm}) = n_m G_m(\hat{\theta}^1_{Mm}, \bar{\theta}^1_M)' W^*_m G_m(\hat{\theta}^1_{Mm}, \bar{\theta}^1_M) - \ln(n_m)(L^*_m - K_{Mm}) \quad (3.21)$$

Here, $G_M(\hat{\theta}^1_m, \bar{\theta}^1_M)$ are the moments for model M evaluated at the estimated values for free parameters $\hat{\theta}^1_{Mm}$ and zero for the restricted parameters, $\bar{\theta}^1_M$; K_{Mm} is the number of free parameters in model M for module m; and n is the number of observations. 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 3.5.

Table 3.5: Share of Modules in which GMM-BIC Criterion Selects Each Model

M	Model Name	Parameter Restrictions	Unweighted Share	Sales Weighted Share
1.	Homothetic	$\gamma_m = 0, \alpha^1_m = 0$	0.15	0.15
2.	Non-Homothetic in Price	$\gamma_m = 0$	0.15	0.10
3.	Non-Homothetic in Quality	$\alpha^1_m = 0$	0.47	0.39
4.	Non-Homothetic in Quality and Price	None	0.26	0.36

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. 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 per cent 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.

In the trade literature, non-homothetic utility functions are either non-homothetic in the demand for quality or non-homothetic in price sensitivity. The J statistics for these two models can be compared directly in Figure 3.7.[33] The J statistic is lower for the model that permits non-homotheticity in the demand for quality in the majority of modules. This indicates that the quality model (Model 2) has a lower GMM-BIC criterion than the price sensitivity model (Model 3) in the majority of those modules for which the most flexible model has the lowest GMM-BIC criterion overall. That is, the model allowing for non-homotheticity in the demand for quality alone explains more of the cross-income variation in consumer behavior than the model that allows for non-homotheticity in price sensitivity alone.

Figure 3.7: Distribution of Log Relative J Statistic for Model 2 Relative to Model 3
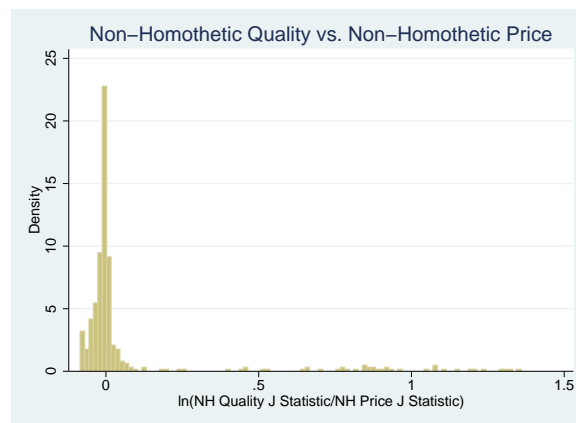
Table 3.6 shows the results of these bilateral model comparisons across all four models. We see that the model that accounts for non-homothetic demand for quality has a lower GMM-BIC criterion in modules representing approximately two-thirds of sales when compared to all three of the other models.

3.5.3 Comparing Utility Across Markets with Relative Price Indexes

As outlined in Section 3.4.2, I use market- and income-specific price indexes as a measure of representative consumer utility from grocery consumption. I compare how the estimated

---

[33]The two elements of the GMM-BIC criterion defined in equation 3.21 that vary by model are the J statistic, $nG_m(\hat{\theta}^1_{Mm}, \bar{\theta}^1_M)'W^*_m G_m(\hat{\theta}^1_{Mm}, \bar{\theta}^1_M)$, and the number of free parameters, K_{Mm}. The models with non-homotheticity in quality or price alone each have one parameter restriction and, therefore, the same number of free parameters.

Table 3.6: Bilateral Model Comparisons: Sales Share of Modules where GMM-BIC(M)<GMM-BIC(N)

| | | Model N | | | |
Model M		1.	2.	3.	4.
1.	Homothetic	-	0.18	0.28	0.32
2.	Non-Homothetic in Price	0.82	-	0.38	0.40
3.	Non-Homothetic in Quality	0.72	0.62	-	0.68
4.	Non-Homothetic in Quality and Price	0.68	0.60	0.32	-

price index for a representative consumer with log income y_k in city c, $\hat{P}(y_k, \mathbb{P}_c)$,[34] varies city-to-city using the following baseline regression model:

$$\ln \hat{P}(y_k, \mathbb{P}_c) = \delta_k + \beta_1 y_c + \beta_2 y_k y_c + \epsilon_{kc}, \tag{3.22}$$

where δ_k is an income-level fixed effect and y_c is log per capita income in city c.

The goal of this analysis is to determine how grocery costs differentially vary across cities for consumers at different income levels. Specifically, equation (3.22) measures how the elasticity of grocery costs in a city with respect to its per capita income varies with household income. The grocery cost price index is calculated using a model that allows for non-homotheticity in the demand for quality, and thus this elasticity will vary with income if the goods available in each city are correlated with the tastes of the incomes of the consumers living there. Suppose, for example, that wealthy cities sell 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 at a faster rate (or increase at a slower rate) than the price index faced by low-income consumers, for those that move from poor to wealthy cities. This is because high-income consumers benefit from the availability and lower prices of the goods that they prefer. This implies that the elasticity of the price index faced by high-income consumers with respect to city income will be lower than the

---

[34]The data includes purchases of a sample of households in each city, as opposed to all households in each city, so I only observe prices (and calculate brand quality) for a subset of the products available in each market. This means that each market- and income-specific price index, $\hat{P}(y_k, \mathbb{P}_c)$, summarizes the utility of consumers from this subset of observed products in each market. The number of products that I observe in each market is increasing in the number of households sampled from that market. The sample household count varies systematically with market population and market income, with correlation coefficients of 0.57 and 0.43, respectively. I control for any potential bias in the results due to these correlations by calculating market-income-specific price indexes over the products purchased by a random sample of 850 households in each market, for each of the 23 markets with 850 or more sample households.

elasticity of the price index faced by low-income consumers with respect to city income. In the above specification, the elasticity of grocery costs with respect to city income is equal to the coefficient on log city income added to the product of log consumer income and by the coefficient on the income interaction term: $\varepsilon_{P_{ck}, y_c} = \beta_1 + \beta_2 y_k$. The results of this regression are presented in the first column of Table 3.7. The estimates for β_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 β_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.

The elasticity of the income-specific price index with respect to city income is shown in Figure 3.8. These elasticities are estimated non-parametrically using 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. Figure 3.8 shows that 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.

In Figure 3.9, I look specifically at the price indexes in San Francisco relative to New Orleans by household income. This chart shows that the variation across these two markets is higher than that predicted by the non-parametric regression coefficients presented in Figure 3.8. Grocery costs are approximately 15 percent higher in San Francisco than those in New Orleans for a consumer with annual income around $15,000, but they are over 60 percent lower in San Francisco for a consumer with annual income around $100,000.

In order to check the robustness of the results using price indexes to measure relative

Table 3.7: City-Income Specific Price Index Regressions

Dependent Variable: Ln(Price Index for Representative Consumer k in City c)				
Ln(Per Capita Income$_c$)	β_1	2.412***	-	2.290*
		[0.996]	-	[1.22]
Ln(Per Capita Income$_c$)	β_2	-0.217**	-	-0.201*
*Ln(Household Income$_k$)		[0.0915]	-	[0.112]
Ln(Population$_c$)	β_3	-	0.28	0.040
		-	[0.190]	[0.232]
Ln(Population$_c$)	β_4	-	-0.027	-0.005
*Ln(Household Income$_k$)		-	[0.018]	[0.021]
Implied Elasticity of Price with respect to City Income ($\beta_1 + \beta_2 y_k$):				
- $y_k = \ln(\$15,000)$		0.325	-	0.357
- $y_k = \ln(\$50,000)$		0.064	-	0.115
- $y_k = \ln(\$100,000)$		-0.086	-	-0.069
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.

expected utility across markets, I measure the relative expected utility between San Francisco and New Orleans using numerical integration. Figure 3.10 shows that this method also shows a 60 percent gap between the utility losses for a low-income household and the utility gains for a high-income household when both get groceries in San Francisco instead of New Orleans.

Figure 3.8: Variation in Elasticity of Grocery Costs with respect to City Income Across Household Income Levels
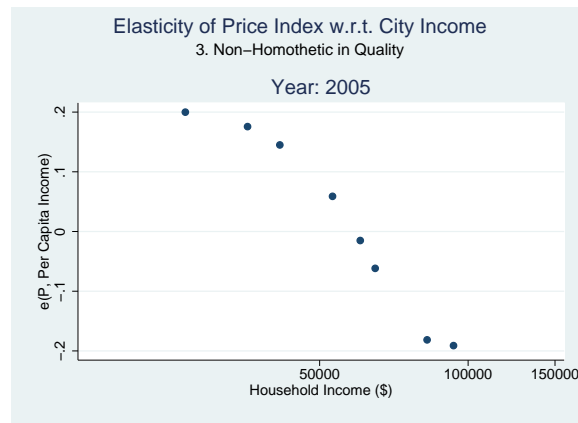
Figure 3.9: Grocery Costs in San Francisco Relative to New Orleans by Household Income
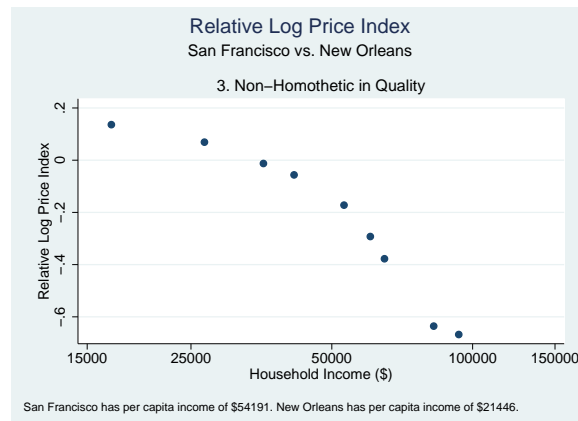
The data presented in Table 3.2 indicates that market income is correlated with market size, that is, wealthier U.S. cities are larger than poorer U.S. cities (the correlation coefficient is 0.47). Therefore, it is possible that the negative coefficient on the market income and household income interaction term in the base line regression (equation 3.22) is due to the fact that grocery costs are lower for high-income households than for low-income

Figure 3.10: Grocery Utility in San Francisco Relative to New Orleans by Household Income
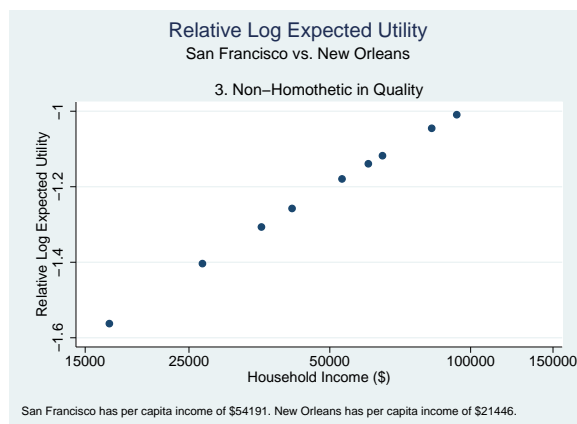
households in larger, as opposed to wealthier, cities. If this were the case, the results above would support a story in which high-income consumers receive more consumption benefits from living in larger cities than low-income consumers, as opposed to the "preference externalities" story in which high-income 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 including log population and log population interacted with log household income in the regressions. The results from these regressions are presented in the second and third columns of Table 3.7. When the log price indexes are regressed against these population variables and household income dummies, 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. These results suggest that, relative to low-income households, high-income households receive higher consumption utility from the grocery bundles available in a wealthier cities than from the grocery bundles available in a poorer cities with the same population size. This pattern is consistent with theories that predict that the composition of demand affects the value of being in a location. Waldfogel (2003)

and Fajgelbaum, Grossman, and Helpman (2011) posit that the gathering of consumers with similar tastes will yield greater consumption benefits to consumers with these tastes than to those with different tastes. The authors refer to this phenomenon as "within-group preference externalities" or "home market effects," respectively. The results above are the first structural estimates of these externalities and provide the first evidence that they are large and economically significant.

I now turn to addressing the extent to which homotheticity biases our 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 showed that there is economically significant variation 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. In Table 3.9 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 rapidly 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.

Table 3.9: Correlation of City-Specific Price Indexes Calculated with Homothetic and Non-Homothetic Models

Household Income	Correlation Between Homothetic Index and Non-Homothetic Index
$16,896	0.83
$26,715	0.88
$35,715	0.94
$41,526	0.96
$53,103	0.97
$60,442	0.92
$64,805	0.83
$82,576	0.21
$93,411	0.03
$146,566	-0.11

Table 3.10 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.[35]

The evidence above shows that non-homothetic preferences yield economically significant variation in living costs: wealthy consumers benefit more from the consumption baskets available in wealthy cities than poor consumers. This variation is particularly relevant for economists who currently use homothetic price indexes to measure real income inequality. Moretti (2008), for example, argues that one should adjust nominal income by location-specific prices when measuring national real income inequality. Since my results

---

[35]See Tables 3.12 and 3.13 in Appendix Section 3.A the the levels and ranks of the homothetic and non-homothetic grocery price indexes in all 23 markets with samples of 850 or more households.

Table 3.10: Cities Ranked According to Grocery Costs Calculated Using Homothetic and Non-Homothetic Models

City Ranks (Least Expensive to Most Expensive)

Six Least Expensive Cities According to Homothetic Model

Market	Homothetic	Non-Homothetic									
		$16,896	$26,715	$35,715	$41,526	$53,103	$60,442	$64,805	$82,576	$93,411	$146,566
Philadelphia	1	1	1	1	1	1	1	1	1	3	3
Washington, DC-Baltimore	2	4	5	2	2	2	2	2	5	11	14
Atlanta	3	7	4	3	3	3	3	3	14	18	21
Detroit	4	11	10	9	7	10	14	19	22	22	22
Columbus	5	2	2	4	4	4	4	6	17	20	20
Tampa	6	3	3	5	5	6	12	13	19	17	9

Six Most Expensive Cities According to Homothetic Model

Market	Homothetic	Non-Homothetic									
		$16,896	$26,715	$35,715	$41,526	$53,103	$60,442	$64,805	$82,576	$93,411	$146,566
Chicago	18	17	18	18	18	17	18	17	12	7	4
San Antonio	19	12	13	16	17	18	17	15	8	6	6
Sacramento	20	19	19	20	21	21	19	16	3	1	1
Minneapolis	21	20	21	21	20	20	21	21	23	23	23
Los Angeles	22	22	22	22	22	22	23	23	20	19	12
San Francisco	23	23	23	23	23	23	22	22	6	4	5

show that high-income consumers face vastly different location-specific prices than low-income consumers, they suggest that economists should adjust nominal income by prices that are both location- and income-specific when measuring national real income inequality. Moretti (2008) finds that the U.S. college wage premium is lower in real terms than in nominal terms because college graduates are concentrated in metropolitan areas where homothetic price indexes are high. The results above indicate that the non-homothetic price indexes faced by high-income college graduates will vary across cities differently to those faced by high school graduates, whose earn lower incomes. If homothetic price indexes are negatively correlated with the non-homothetic price index for high-income consumers, as the results in Table 3.9 suggest is the case for the groceries, then Moretti (2008) will tend to underestimate the real income of college graduates relative to high school graduates. Measuring the size of this bias will require cost-of-living indexes that account for non-homotheticity in demand for all products, services and housing and is left to future research.

3.6   Conclusion

There is growing interest in the role of non-homothetic preferences and cross-market income differences in determining production patterns in macroeconomics and trade. 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. I show that this is indeed the case: high-income households face greater consumption gains from moving to high per capita income markets than low-income households.

I measure the extent of this variation using a multi-sector utility function that allows for non-homothetic demand for quality. This model is based on a framework that embeds the two forms of non-homotheticity that are most often used to model how demand varies with income. This framework yields estimating equations in which the extent to which the elasticity of demand with respect to quality varies with income is separately identifiable from the extent to which the elasticity of demand with respect to price varies with income. While previous work has used product unit values to proxy for quality when testing for

evidence of non-homotheticity of demand with respect to quality, I estimate product quality non-parametrically. This permits the separate identification of parameters that govern the non-homotheticity of demand with respect to estimated quality and those that govern the non-homotheticity of demand with respect to observed prices. The data shows more widespread evidence that preferences are non-homothetic with respect to quality than with respect to price. In most of the grocery sectors over which demand is estimated, the GMM-BIC model selection test picks the model that allows for non-homothetic demand for quality but not for non-homothetic price sensitivity. Further research is required to test whether this is the case for a wider set of products.

I show how non-homotheticity impacts our estimates of relative utility across U.S. cities, *i.e.* domestic markets separated by geography. Specifically, I find that the utility from consumption in a city relative to another city with half the per capita income is 40 percent higher for a high-income household than for a low-income household. The framework presented could also be used to measure the cross-income variation in the relative utility of consumers across international markets or across markets separated by time or trade liberalization events. Such analysis would improve our understanding of the implications of non-homothetic preferences for the measurement of purchasing power parity, inflation, and gains from trade. While the estimates presented here are generated using household-level consumption data, it is conceivable that simulation techniques pioneered in the IO field could be used to identify the parameters of the model using aggregate market- or country- level data. If so, the framework presented above could be used to measure how relative welfare varies with income across international markets. Future work will investigate this possibility and, if successful, will use the model presented above to measure income specific purchasing power parity deflators. Additionally, the model has the potential to be used to show how the gains from trade for consumers vary with income over recent decades of liberalization.

3.A   Market Regions and Price Indexes

Table 3.11: Regional Categorizations for Sample Markets

| Market Code	Market	Region	Neighboring Region
1	Des Moines	MW	SC
2	Little Rock	SE	SC
3	Omaha	NW	MW
4	Syracuse	NE	MW
5	Albany	NE	MW
6	Birmingham	SE	SC
7	Richmond	NE	SE
8	Louisville	MW	SE
9	Grand Rapids	MW	NE
10	Jacksonville	SE	SC
11	Memphis	SE	SC
12	Raleigh-Durham	SE	SC
13	Nashville	SE	SC
14	Salt Lake City	SW	SC
15	Charlotte	SE	SC
16	Columbus	MW	NE
17	San Antonio	SW	SC
18	Indianapolis	SE	MW
19	Orlando	SE	SC
20	Milwaukee	MW	NW
21	Hartford-New Haven	NE	MW
22	Kansas City	MW	SE
23	Sacramento	SW	NW
24	New Orleans-Mobile	SE	SC
25	Oklahoma City-Tulsa	SC	MW
26	Cincinnati	MW	NE
27	Portland, Or	NW	SW
28	Buffalo-Rochester	NE	MW
29	Pittsburgh	MW	NE
30	Tampa	SE	SC
31	Denver	SW	NW
32	St. Louis	MW	NE
33	San Diego	SW	NW
34	Cleveland	MW	NE
35	Minneapolis	MW	NW
36	Phoenix	SW	SC
37	Seattle	NW	SW
38	Miami	SE	SC
39	Atlanta	SE	SC
40	Houston	SC	SC
41	Dallas	SC	SW
42	Detroit	MW	NE
43	Boston	NE	MW
44	Philadelphia	NE	MW
45	San Francisco	NW	SW
46	Washington, DC-Baltimore	NE	SE
47	Chicago	MW	NE
48	Los Angeles	SW	SC
49	New York	NE	MW

Table 3.12: City-Specific Price Indexes Calculated Using Homothetic and Non-Homothetic Models

| Market | Price Index (Relative to the Homothetic Index for New York) | | | | | | | | | | |
| | Homothetic | Non-Homothetic | | | | | | | | | |
		$16,896	$26,715	$35,715	$41,526	$53,103	$60,442	$64,805	$82,576	$93,411	$146,566
Philadelphia	0.83	0.77	0.81	0.83	0.84	0.84	0.84	0.84	0.84	0.84	0.82
Washington, DC-Baltimore	0.83	0.86	0.87	0.86	0.85	0.85	0.87	0.88	0.96	0.99	1.03
Atlanta	0.87	0.88	0.85	0.86	0.86	0.88	0.89	0.91	1.00	1.05	1.16
Detroit	0.93	0.98	0.98	0.95	0.95	0.98	1.02	1.06	1.18	1.21	1.18
Columbus	0.94	0.77	0.83	0.88	0.90	0.92	0.93	0.94	1.03	1.07	1.15
Tampa	0.95	0.79	0.83	0.88	0.91	0.96	1.00	1.02	1.07	1.05	0.97
Houston	0.97	0.89	0.91	0.94	0.95	0.97	0.98	0.98	0.98	0.97	0.96
St. Louis	0.98	1.12	1.05	1.00	0.99	0.97	0.97	0.98	1.02	1.04	1.07
Phoenix	0.98	0.94	0.95	0.96	0.97	0.96	0.95	0.94	0.93	0.94	0.98
Dallas	0.98	0.88	0.90	0.93	0.94	0.95	0.96	0.96	1.00	1.03	1.10
New York	1.00	0.88	0.99	1.04	1.04	1.00	0.95	0.93	0.84	0.82	0.81
Miami	1.01	0.94	0.94	0.95	0.96	0.98	0.98	0.98	0.97	0.98	1.03
Boston	1.02	1.04	1.06	1.05	1.04	1.02	1.01	1.00	1.02	1.03	1.05
Charlotte	1.02	1.06	1.01	0.99	0.99	1.01	1.03	1.05	1.05	1.02	0.98
Denver	1.02	1.04	1.02	1.02	1.03	1.04	1.03	1.03	1.00	0.98	0.96
Buffalo-Rochester	1.03	1.27	1.19	1.11	1.08	1.07	1.08	1.09	1.13	1.13	1.08
Seattle	1.03	1.05	1.03	1.03	1.02	1.01	1.00	0.99	0.99	1.01	1.05
Chicago	1.05	1.09	1.10	1.08	1.07	1.05	1.04	1.04	1.00	0.97	0.90
San Antonio	1.08	1.00	1.02	1.05	1.06	1.06	1.04	1.03	0.98	0.96	0.95
Sacramento	1.10	1.21	1.15	1.13	1.13	1.11	1.07	1.04	0.86	0.80	0.74
Minneapolis	1.13	1.23	1.20	1.15	1.12	1.09	1.09	1.10	1.19	1.24	1.26
Los Angeles	1.15	1.29	1.23	1.18	1.17	1.17	1.17	1.16	1.10	1.06	1.01
San Francisco	1.22	1.33	1.29	1.26	1.24	1.20	1.14	1.10	0.96	0.93	0.93
Correlation with Homothetic Index		0.83	0.88	0.94	0.96	0.97	0.92	0.83	0.21	0.03	-0.11

Table 3.13: Cities Ranked According to Grocery Costs Calculated Using Homothetic and Non-Homothetic Models

Market	Homothetic	City Ranks (Least Expensive to Most Expensive) Non-Homothetic									
		$16,896	$26,715	$35,715	$41,526	$53,103	$60,442	$64,805	$82,576	$93,411	$146,566
Philadelphia	1	1	1	1	1	1	1	1	1	3	3
Washington, DC-Baltimore	2	4	5	2	2	2	2	2	5	11	14
Atlanta	3	7	4	3	3	3	3	3	14	18	21
Detroit	4	11	10	9	7	10	14	19	22	22	22
Columbus	5	2	2	4	4	4	4	6	17	20	20
Tampa	6	3	3	5	5	6	12	13	19	17	9
Houston	7	8	7	7	8	8	10	10	9	8	8
St. Louis	8	18	16	12	12	9	8	9	16	16	17
Phoenix	9	9	9	10	10	7	5	5	4	5	11
Dallas	10	5	6	6	6	5	7	7	13	14	19
New York	11	6	11	15	15	12	6	4	2	2	2
Miami	12	10	8	8	9	11	9	8	7	10	13
Boston	13	14	17	17	16	15	13	12	15	15	15
Charlotte	14	16	12	11	11	14	16	18	18	13	10
Denver	15	13	14	13	14	16	15	14	11	9	7
Buffalo-Rochester	16	21	20	19	19	19	20	20	21	21	18
Seattle	17	15	15	14	13	13	11	11	10	12	16
Chicago	18	17	18	18	18	17	18	17	12	7	4
San Antonio	19	12	13	16	17	18	17	15	8	6	6
Sacramento	20	19	19	20	21	20	19	16	3	1	6
Minneapolis	21	20	21	21	20	20	21	21	23	23	23
Los Angeles	22	22	22	22	22	22	23	23	20	19	12
San Francisco	23	23	23	23	23	23	22	22	6	4	5

Figure 3.11: Income-Specific Food Expenditure Shares Across Markets
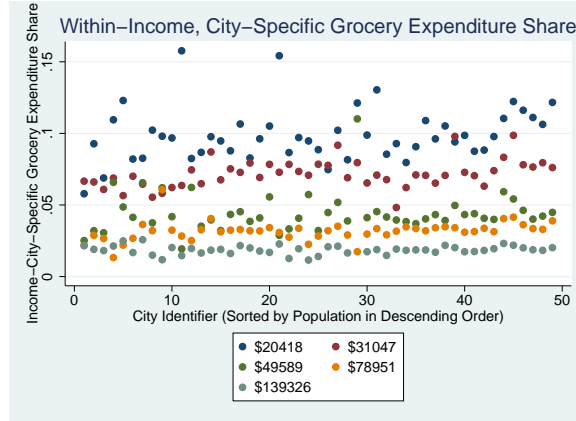
3.B Non-Homotheticity Condition

Recall from equation 3.1 that consumers select grocery consumption quantities, $\mathbb{Q} = \{\{q_{mg}\}_{g \in \mathbf{G_m}}\}_{m \in \mathbf{M}}$, and non-grocery expenditure, Z, by maximizing:

$$f(U_{iG}(\mathbb{Q},Z), Z) \quad \text{subject to} \quad \sum_{m \in \mathbf{M}} \sum_{g \in \mathbf{G_m}} p_{mg} q_{mg} + Z \leq Y_i, \ q_{mg} \geq 0 \ \forall \ mg \in \mathbf{G} \tag{3.23}$$

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

$$\max_{\mathbb{Q},Z} U_{iG}(\mathbb{Q}, Z) = \left\{ \sum_{m \in M} \left[\sum_{g \in G_m} q_{mg} \exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img}) \right]^{\frac{\sigma(Z)-1}{\sigma(Z)}} \right\}^{\frac{\sigma(Z)}{\sigma(Z)-1}}$$

$$\text{subject to} \quad \sum_{m \in \mathbf{M}} \sum_{g \in \mathbf{G_m}} p_{mg} q_{mg} \leq Y_i - Z, \ q_{mg} \geq 0 \ \forall \ mg \in \mathbf{G} \tag{3.24}$$

where $b_{mg}(Z) = \beta_{mg}(1 + \gamma_m z)$, $\mu_m(Z) = \frac{1}{\alpha_m^0 + \alpha_1^1 z}$, and $\sigma(Z) = 1 + \alpha^0 + \alpha^1 z$ for $z = \ln Z$. The solution to this problem for a consumer i with idiosyncratic utility draws ε_i is solved in the paper. Equations (3.9), (3.12), and (3.13) define the optimal grocery bundle, $\mathbb{Q}^*(Z) = \left\{ \left\{ q_{mg}^*(Z) \right\}_{g \in \mathbf{G},} \right\}_{m \in \mathbf{M}}$:

$$q_{img}^*(Z) = \begin{cases} (Y_i - Z)\dfrac{\left[\frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{ig})}{p_{mg}}\right]^{\sigma(Z)-1}}{P_i(Z)^{1-\sigma(Z)}}/p_{mg} & \text{if } g = \arg\max_{g\in\mathbf{G_m}} \dfrac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{img})}{p_{mg}} \\ 0 & \text{otherwise} \end{cases}$$

where

$$P_i(Z) = \left[\left(\sum_{m\in\mathbf{M}} \max_{g\in\mathbf{G_m}} \frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{ig})}{p_{mg}}\right)^{\sigma(Z)-1}\right]^{\frac{1}{1-\sigma(Z)}}$$

Plugging this solution into $U_{iG}(\mathbb{Q}, Z)$ yields the consumer's indirect utility from grocery consumption, conditional on their non-grocery expenditure:

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

$$= \left\{\sum_{m\in M}\left[\left((Y_i-Z)\frac{\left(\frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{ig})}{p_{mg}}\right)^{\sigma(Z)}}{P_i(Z)^{1-\sigma(Z)}}\right)\mathbb{I}\left[g=\arg\max_{g\in\mathbf{G_m}}\frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]\right]^{\frac{\sigma(Z)-1}{\sigma(Z)}}\right\}^{\frac{\sigma(Z)}{\sigma(Z)-1}}$$

$$= \frac{Y_i-Z}{P_i(Z)^{1-\sigma(Z)}}\left\{\sum_{m\in M}\left[\left(\frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{ig})}{p_{mg}}\right)^{\sigma(Z)}\mathbb{I}\left[g=\arg\max_{g\in\mathbf{G_m}}\frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]\right]^{\frac{\sigma(Z)-1}{\sigma(Z)}}\right\}^{\frac{\sigma(Z)}{\sigma(Z)-1}}$$

$$= \frac{Y_i-Z}{P_i(Z)^{1-\sigma(Z)}}\left\{\sum_{m\in M}\left(\max_{g\in\mathbf{G_m}}\frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{ig})}{p_{mg}}\right)^{\sigma(Z)-1}\right\}^{\frac{\sigma(Z)}{\sigma(Z)-1}}$$

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

$$= \frac{Y_i-Z}{P_i(Z)} \tag{3.25}$$

We can now express problem (3.23) to be a choice over one variable, Z:

$$\max_Z f(\tilde{U}_{iG}(Z), Z) \tag{3.26}$$

The first order condition to the utility maximization problem defined in problem (3.26) with respect to Z is:

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

If we assume that $f(U_{iG}(\mathbb{Q}, Z), Z)$ is additive in grocery utility and non-grocery utility, i.e. $f(U_{iG}(\mathbb{Q}, Z), Z) = U_{iG}(\mathbb{Q}, Z) + Z$, the formula above simplifies to:

$$\frac{\partial \tilde{U}_i(Z)}{\partial Z} + 1 = 0$$

Substituting the maximized grocery expenditure conditional on Z, $\tilde{U}_{iG}(Z)$, from equation (3.25) 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, ε_i, and model parameters:

$$\frac{\partial}{\partial Z}\left[\frac{Y_i - Z}{P_i(Z)}\right] + 1 = 0$$

We can simplify this expression to get income as a function of prices and outside good expenditure:

$$\frac{-P_i(Z) - P_i'(Z)(Y_i - Z)}{P_i(Z)^2} + 1 = 0$$

$$P_i(Z) + P_i'(Z)(Y_i - Z) = P_i(Z)^2$$

$$Y_i = \frac{P_i(Z)^2 - P_i(Z)}{P_i'(Z)}$$

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 is such that:

$$\frac{\partial}{\partial Z}\frac{P_i(Z)^2 - P_i(Z)}{P_i'(Z)} > 0$$

3.C   Connection to Nested CES Utility Function

Consider the related utility function for some consumer i who is representative of all consumers with outside good expenditure Z:

$$U_i = \left\{\sum_{m \in M}\left[\sum_{g \in G_m}[q_{mg}\exp(b_{mg}(Z))]^{\rho_m(Z)}\right]^{\frac{\rho(Z)}{\rho_m(Z)}}\right\}^{\frac{1}{\rho(Z)}}$$

where $b_{mg}(Z) = \beta_{mg}(1 + \gamma_m z_i)$, $\rho_m(Z) = \frac{1 - \sigma_m(Z)}{\sigma_m(Z)}$ for $\sigma_m(Z) = 1 + \alpha_m^0 + \alpha_m^1 z$, $\rho(Z) = \frac{1 - \sigma(Z)}{\sigma(Z)}$ for $\sigma(Z) = 1 + \alpha^0 + \alpha^1 z$, and $z = \ln(Z)$. Suppose that this representative consumer faces the same prices $\mathbb{P}$ and has the same outside good expenditure Z as a group of consumers with the CES-nested log-logit utility defined in equation (3.2). A simple extension of the Anderson, de Palma, and Thisse (1987) result implies that the representative consumer's within-module expenditure shares will be identical to the within-module market shares of a group of consumers with the same outside good expenditure. Below, I show that the same is true for the representative consumer's between-module expenditure shares. For the sake of comparison, consider the representative consumer's log expenditure in module m relative to module $\bar{m}$:

$$\ln s_{im} - \ln s_{i\bar{m}} = -(\alpha^0 + \alpha^1 z)\left[\ln P_m(z, \mathbb{P}_m) - \ln P_{\bar{m}}(z, \mathbb{P}_{\bar{\mathbf{m}}})\right] \tag{3.27}$$

where $P_m(z, \mathbb{P}_m)$ is a CES price index defined as:

$$P_m(z, \mathbb{P}_m) = \left[\sum_{g \in \mathbf{G_m}}\left(\frac{p_{mg}}{\exp(\beta_{mg}(1 + \gamma_m z))}\right)^{-(\alpha_m^0 + \alpha_m^1 z)}\right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 z)}} \tag{3.28}$$

The representative consumer's relative log expenditure share is inversely proportional to the difference in the quality-adjusted price levels in the modules. However, the idiosyncratic consumer's relative log expenditure share is proportional to the difference in price-adjusted quality levels for each module. These relative log expenditure shares are equivalent because the quality-adjusted price levels defined in equation (3.28) are equal to the inverse of the price-adjusted quality levels defined in equation (3.37). The income-specific quality-adjusted price of a product g in module m is equal to the inverse of the income-specific price-adjusted quality of the same product:

$$
\begin{aligned}
P_m(z, \mathbb{P}_m) &= \left[\sum_{g \in \mathbf{G_m}} \left(\frac{p_{mg}}{\exp(\beta_{mg}(1 + \gamma_m z))} \right)^{-(\alpha_m^0 + \alpha_m^1 z)} \right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 z)}} \\
&= \left[\sum_{g \in \mathbf{G_m}} \left(\left(\frac{\exp(\beta_{mg}(1 + \gamma_m z))}{p_{mg}} \right)^{-1} \right)^{-(\alpha_m^0 + \alpha_m^1 z)} \right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 z)}} \\
&= \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\beta_{mg}(1 + \gamma_m z))}{p_{mg}} \right)^{(\alpha_m^0 + \alpha_m^1 z)} \right]^{\frac{1}{-(\alpha_m^0 + \alpha_m^1 z)}} \\
&= \left\{ \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\beta_{mg}(1 + \gamma_m z))}{p_{mg}} \right)^{(\alpha_m^0 + \alpha_m^1 z)} \right]^{\frac{1}{(\alpha_m^0 + \alpha_m^1 z)}} \right\}^{-1} \\
&= [V_m(z, \mathbb{P}_m)]^{-1}
\end{aligned}
$$

3.D   Solving for Optimal Between-Module Expenditure Shares

Consumer i, spending Z on the outside good, chooses how to allocate expenditures between modules by selecting $w_1, ..., w_M$ to maximize

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

subject to

$$
\sum_{m \in \mathbf{M}} w_m \leq W_i
$$

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}:

$$
\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}} = \frac{1}{A_{im}} \tag{3.29}
$$

The within-module allocation decision now simplifies to:

$$
\mathbf{w_i^*} = (w_{i1}^*, ..., w_{iM}^*) = \underset{\sum_{m \in \mathbf{M}} w_m \leq W_i}{\arg \max} \left\{ \sum_{m \in M} \left[\frac{w_m}{A_{im}} \right]^{\frac{\sigma(Z)-1}{\sigma(Z)}} \right\}^{\frac{\sigma(Z)}{\sigma(Z)-1}} \tag{3.30}
$$

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 with respect to expenditure in each module m. The first order condition for each module m is:

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

where λ 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':

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

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

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

Imposing the budget constraint, $\sum_{m\in\mathbf{M}}w_{m'} = \sum_{m\in\mathbf{M}}w_m \leq W_i$, yields an expression for w_m in terms of total expenditure, W_i, and an index of the A_{im} terms:

$$W = \sum_{m'\in\mathbf{M}}w_{m'}$$

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

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

The solution to problem (3.30) is, therefore,

$$\mathbf{w_i^*} = (w_{i1}^*,...,w_{iM}^*) \quad \text{where} \quad w_{im}^* = \frac{A_{im}^{1-\sigma(Z)}}{P_i^{1-\sigma(Z)}}W_i \;\; \forall m\in\mathbf{M}$$

where P_i is a CES price index over A_{im} for all modules $m\in\mathbf{M}$ defined as:

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

Substituting from equation (3.29) for A_{img} yields consumer i's optimal module expenditure vector,

$\mathbf{w_i^*}$, as a function of total grocery expenditures, prices, and model parameters:

$$\mathbf{w_i^*} = (w_{i1}^*, ..., w_{iM}^*) \quad \text{where} \quad w_{im}^* = W_i \frac{\left[\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]^{\sigma(Z)-1}}{P_i^{1-\sigma(Z)}}$$

$$P_i = \left[\sum_{m \in \mathbf{M}} \left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)^{\sigma(Z)-1}\right]^{\frac{1}{1-\sigma(Z)}}$$

3.E Deriving Within-Module Market Expenditure Shares

Equation (3.9) states that:

$$\mathbb{Q}_{im}^*(w_m) = (q_{im1}^*(w_m), \ldots, q_{imG_m}^*(w_m)) \text{ where } q_{img}^*(w_m) = \begin{cases} w_m/p_{mg} & \text{if } g = \arg\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}} \\ 0 & \text{otherwise} \end{cases}$$

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:

$$q_{img}^*(w_m) = \frac{w_m}{p_{mg}} \mathbb{I}\left[g = \arg\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]$$

We can use this definition to derive consumer i's expenditure on product g in module m:

$$w_{img}(w_m) = p_{mg}q_{img}^*(w_m) = w_m \mathbb{I}\left[g = \arg\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]$$

Dividing through by w_m yields the consumer's expenditure share on product g in module m:

$$s_{img|m} = \mathbb{I}\left[g = \arg\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]$$

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

$$\begin{aligned} \mathbb{E}_\varepsilon[s_{img|m}] &= \mathbb{E}_\varepsilon\left[\mathbb{I}\left[g = \arg\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]\right] \\ &= Pr\left[\frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}} \geq \frac{\exp(b_{mg'}(Z) + \mu_m(Z)\varepsilon_{img'})}{p_{mg'}}, \ \ \forall g' \in \mathbf{G_m}\right] \\ &= Pr\left[\varepsilon_{img} - \varepsilon_{img'} \geq \frac{b_{mg}(Z) - b_{mg'}(Z) - (\ln p_{mg} - \ln p_{mg'})}{\mu_m(Z)}, \ \ \forall g' \in \mathbf{G_m}\right] \\ &= \frac{\exp \frac{b_{mg}(Z) - \ln p_{mg}}{\mu_m(Z)}}{\sum_{g' \in \mathbf{G_m}}\left(\exp \frac{b_{mg'}(Z) - \ln p_{g'}}{\mu_m(Z)}\right)} \end{aligned}$$

The final equality holds because the idiosyncratic utilities, ε_{im}, are iid draws from a type I extreme value distribution. Imposing the parametric forms for $b_{mg}(Z) = \beta_{mg}(1 + \gamma_m z_i)$ and $\mu_m(Z) = (\alpha_m^0 + \alpha_m^1 z_i)^{-1}$ from equations (3.5) and (3.6), respectively, ensures that the consumer's expected expenditure share is

common with other consumers with the same income that face the same product prices:

$$\mathbb{E}_\varepsilon[s_{img|m}] = \frac{\exp[(\alpha_m^0 + \alpha_m^1 z_i)(\beta_{mg}(1 + \gamma_m z_i) - \ln p_{mg})]}{\sum\limits_{g' \in \mathbf{G_m}} \left(\exp[(\alpha_m^0 + \alpha_m^1 z_i)(\beta_{mg'}(1 + \gamma_m z_i) - \ln p_{mg'})]\right)}$$

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, $Z_i = \exp(z_i)$, 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 $s_{mg|m}(Z_i, \mathbb{P}_m)$. $s_{mg|m}(Z_i, \mathbb{P}_m)$ is the share of expenditure that a group of households with the outside good expenditure, Z_i, and facing a common vector of module prices, $\mathbb{P}_m$:

$$s_{mg|m}(Z_i, \mathbb{P}_m) = \mathbb{E}_\varepsilon[s_{img|m}] = \frac{\exp[(\alpha_m^0 + \alpha_m^1 z_i)(\beta_{mg}(1 + \gamma_m z_i) - \ln p_{mg})]}{\sum\limits_{g' \in \mathbf{G_m}} \left(\exp[(\alpha_m^0 + \alpha_m^1 z_i)(\beta_{mg'}(1 + \gamma_m z_i) - \ln p_{mg'})]\right)}$$

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 $\bar{g}_m$[36]:

$$\frac{s_{mg|m}(Z_i, \mathbb{P}_m)}{s_{m\bar{g}|m}(Z_i, \mathbb{P}_m)} = \frac{\exp[(\alpha_m^0 + \alpha_m^1 z_i)(\beta_{mg}(1 + \gamma_m z_i) - \ln p_{mg})]}{\exp[(\alpha_m^0 + \alpha_m^1 z_i)(\beta_{m\bar{g}}(1 + \gamma_m z_i) - \ln p_{m\bar{g}})]}$$

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

$$\ln(s_{mg|m}(Z_i, \mathbb{P}_m)) - \ln(s_{m\bar{g}|m}(Z_i, \mathbb{P}_m)) = (\alpha_m^0 + \alpha_m^1 z_i)\left[(\beta_{mg} - \beta_{m\bar{g}})(1 + \gamma_m z_i) - (\ln p_{mg} - \ln p_{m\bar{g}})\right] \quad (3.31)$$

Equation (3.31) 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 in module m relative to product $\bar{g}_m$ in the same module m in terms of parameters α_m, γ_m, and $(\beta_{mg} - \beta_{m\bar{g}_m})$. This equation will be use to define moments for each product $g \neq \bar{g}_m$ in each module m, and these moments will be used to estimate all of the α_m and γ_m parameters, as well as each β_{mg} parameter relative to $\beta_{m\bar{g}_m}$, i.e. $\{\beta_{mg} - \beta_{m\bar{g}_m}\}_{g \in \mathbf{G_m}}$.

3.F   Deriving Between-Module Relative Market Expenditure Shares

I now want to generate a similar estimation equation that can be used to identify α^0, α^1, and $\{\beta_{\bar{g}_m}\}_{g \in \mathbf{G_m}}$ using data on module-level income-specific market shares. Equations (3.12) and (3.13) together characterize the optimal cross-module expenditure allocation for consumer i conditional on this consumer's idiosyncratic

---

[36]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.

utility draws for each product in each module. These equations are:

$$\mathbf{w_i^*} = (w_{i1}^*, ..., w_{iM}^*) \text{ where } w_{im}^* = W_i \frac{\left[\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]^{\sigma(Z)-1}}{P_i^{1-\sigma(Z)}}$$

$$P_i = \left[\sum_{m \in \mathbf{M}} \left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)^{\sigma(Z)-1}\right]^{\frac{1}{1-\sigma(Z)}}$$

I generate consumer i's optimal module m expenditure share, s_{im}, by dividing their optimal module expenditure, w_{im}^*, by total grocery expenditure W_i :

$$s_{im} = \frac{w_{im}^*}{W_i} = \frac{\left[\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]^{\sigma(Z)-1}}{P_i^{1-\sigma(Z)}}$$

When deriving the within-module relative market share, equation (3.31) above, I take the expectation of the consumer's expected product expenditure share over the idiosyncratic errors, $\mathbb{E}_\varepsilon[s_{img|m}]$, 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.[37]

I now generate the relative module market shares. As discussed above, I first divide consumer i's

---

[37]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:

$$\begin{aligned}\ln(s_{mg|m}(Y_i, \mathbb{P}_m)) - \ln(s_{m\bar{g}|m}(Y_i, \mathbb{P}_m)) &= \ln\left[\mathbb{E}_\varepsilon[s_{img|m}]/\mathbb{E}_\varepsilon[s_{i\bar{m}g|\bar{m}}]\right] \\ &= \mathbb{E}_\varepsilon\left[\ln(s_{img|m}) - \ln(s_{im\bar{g}|m})\right] \\ &= (\alpha_m^0 + \alpha_m^1 y_i)\left[(\beta_{mg} - \beta_{m\bar{g}})(1 + \gamma_m y_i) - (\ln p_{mg} - \ln p_{m\bar{g}})\right]\end{aligned}$$

I derive the expression for the income-specific market share of product g, $s_{mg|m}(Y_i, \mathbb{P}_m) = \mathbb{E}_\varepsilon[s_{img|m}]$, before taking logs and differencing to generate the estimation equation (3.31), as it demonstrates the relationship between the term on the left-hand side of this equation, $\ln(s_{mg|m}(Y_i, \mathbb{P}_m)) - \ln(s_{m\bar{g}|m}(Y_i, \mathbb{P}_m))$, and its value in the data: the difference between the log of the expenditure consumers earning Y_i 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 income-specific market shares on products g and $\bar{g}$.

module expenditure share, s_{im}, by his/her expenditure share in some fixed base module $\bar{m}$:

$$\frac{s_{im}}{s_{i\bar{m}}} = \frac{\left[\max_{g \in \mathbf{G_m}} \dfrac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right]^{\sigma(Z)-1}}{\left[\max_{g \in \mathbf{G_{\bar{m}}}} \dfrac{\exp(b_{,\bar{m}g}(Z) + \mu_{\bar{m}}(Z)\varepsilon_{i\bar{m}g})}{p_{\bar{m}g}}\right]^{\sigma(Z)-1}}$$

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:

$$\ln s_{im} - \ln s_{i\bar{m}} = (\sigma(Z)-1) \ln \left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}} \right) - (\sigma(Z)-1) \ln \left(\max_{g \in \mathbf{G_{\bar{m}}}} \frac{\exp(b_{\bar{m}g}(Z) + \mu_m(Z)\varepsilon_{i\bar{m}g})}{p_{\bar{m}g}} \right),$$

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:

$$\mathbb{E}_\varepsilon\left[\ln s_{im} - \ln s_{i\bar{m}}\right] = (\sigma(Z)-1) \left\{ \mathbb{E}_\varepsilon\left[\ln\left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)\right] - \mathbb{E}_\varepsilon\left[\ln\left(\max_{g \in \mathbf{G_{\bar{m}}}} \frac{\exp(b_{\bar{m}g}(Z)+\mu_m(Z)\varepsilon_{i\bar{m}g})}{p_{\bar{m}g}}\right)\right] \right\}$$

$$(3.32)$$

Consider the two expectation terms in equation (3.32). Both take the same form, and thus I only solve for the first expectation:

$$\mathbb{E}_\varepsilon\left[\ln\left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)\right] \qquad (3.33)$$

The expectation term defined in equation (3.33) 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:

$$\mathbb{E}_\varepsilon\left[\ln\left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)\right] = \mathbb{E}_\varepsilon\left[\max_{g \in \mathbf{G_m}} \ln\left(\frac{\exp(b_{mg}(Z)+\mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)\right]$$

$$= \mathbb{E}_\varepsilon\left[\max_{g \in \mathbf{G_m}} b_{mg}(Z) - \ln p_{mg} + \mu_m(Z)\varepsilon_{img}\right]$$

$$= \mu_{im}\mathbb{E}_\varepsilon\left[\max_{g \in \mathbf{G_m}} (b_{mg}(Z) - \ln p_{mg})/\mu_m(Z) + \varepsilon_{img}\right] \quad (3.34)$$

De Palma and Kilani (2007) show that, for an additive random utility model with $u_i = \nu_i + \varepsilon_i$, $i = 1, \dots, n$ and $\varepsilon_i \overset{\text{iid}}{\sim} F(x)$ a continuous CDF with finite expectation, the expected maximum utility is:

$$\mathbb{E}_\varepsilon[\max_i \nu_i + \varepsilon_i] = \int_{-\infty}^{\infty} z\, d\phi(z) \text{ where } \phi(z) = Pr[\max_k \nu_k \le z] = \prod_{k=1}^{n} F(z - \nu_k)$$

Since the expectation in equation (3.34) takes the form $\mathbb{E}_\varepsilon[\max_g \nu_{img} + \varepsilon_{img}]$, with $\nu_{img} = (b_{mg}(Z) - \ln p_{mg})/\mu_m(Z)$, and since I have assumed that $\varepsilon_{img} \overset{\text{iid}}{\sim} F(x)$ 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:

$$\mathbb{E}_\varepsilon \left[\max_{g \in \mathbf{G_m}} v_g + \varepsilon_g \right] = \int_{-\infty}^{\infty} z \, d\phi(z)$$

$$= \int_{-\infty}^{\infty} z \, d \left[\prod_{g=1}^{G_m} \exp(-\exp(v_g - z)) \right]$$

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

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

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 $dz = -(1/x)dx$:

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

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

$$= \int_{\infty}^{0} (V - \ln x) \exp(-x) \, x(-1/x)dx$$

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

$$= V$$

Since we have defined $\nu_{img} = (b_{mg}(Z) - \ln p_{mg})/\mu_m(Z)$ 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 (3.33):

$$\mathbb{E}_\varepsilon \left[\ln \left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}} \right) \right] = \mu_m(Z) \ln \left[\sum_{g \in \mathbf{G_m}} \exp((b_{mg}(Z) - \ln p_{mg})/\mu_m(Z)) \right]$$

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

$$= \ln \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(b_{mg}(Z))}{p_{mg}} \right)^{\frac{1}{\mu_m(Z)}} \right]^{\mu_m(Z)} \quad (3.35)$$

Plugging this result back into equation (3.32) yields the expected relative module expenditure share for consumer i in terms of product prices and model parameters:

$$\mathbb{E}_\varepsilon \left[\ln s_{im} - \ln s_{i\bar{m}} \right] = (\sigma(Z) - 1) \left\{ \mathbb{E}_\varepsilon \left[\ln \left(\max_{g \in \mathbf{G_m}} \frac{\exp(b_{mg}(Z) + \mu_m(Z)\varepsilon_{img})}{p_{mg}} \right) \right] - \mathbb{E}_\varepsilon \left[\ln \left(\max_{g \in \mathbf{G_{\bar{m}}}} \frac{\exp(b_{\bar{m}g}(Z) + \mu_{\bar{m}}(Z)\varepsilon_{i\bar{m}g})}{p_{\bar{m}g}} \right) \right] \right\}$$

$$= (\sigma(Z) - 1) \left\{ \ln \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(b_{mg}(Z))}{p_{mg}} \right)^{\frac{1}{\mu_m(Z)}} \right]^{\mu_m(Z)} - \ln \left[\sum_{g \in \mathbf{G_{\bar{m}}}} \left(\frac{\exp(b_{\bar{m}g}(Z))}{p_{\bar{m}g}} \right)^{\frac{1}{\mu_{\bar{m}}(Z)}} \right]^{\mu_{\bar{m}}(Z)} \right\}$$

This function only varies by consumer through their outside good expenditure. All consumers with the same outside good expenditure, $Z_i = \exp(z_i) - 1$, that face the same prices, $\mathbb{P}$, will have the same expected

relative module expenditure share:

$$\mathbb{E}_{\varepsilon}\left[\ln s_{im} - \ln s_{i\bar{m}}\right] \quad = \quad -(\alpha^0 + \alpha^1 z_i)\left[\ln V_m(z_i, \mathbb{P}_m) - \ln V_{\bar{m}}(z_i, \mathbb{P}_{\bar{\mathbf{m}}})\right] \tag{3.36}$$

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

$$V_m(z_i, \mathbb{P}_m) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\beta_{mg}(1 + \gamma_m z_i))}{p_{mg}}\right)^{\alpha_m^0 + \alpha_m^1 y_i}\right]^{\frac{1}{\alpha_m^0 + \alpha_m^1 y_i}} \tag{3.37}$$

Equations (3.36) and (3.37) together define the expected relative module expenditure share of a set of households with income Y_i that face prices $\mathbb{P}_m$ and $\mathbb{P}_{\bar{m}}$ in terms of parameters α^0, α^1, as well as α_m, γ_m, β_{mg} for all $g \in G_m$, and $\alpha_{\bar{m}}$, $\gamma_{\bar{m}}$, $\beta_{\bar{m}g}$ for all $g \in G_{\bar{m}}$.

Bibliography

ADAMSON, D. W., D. E. CLARK, AND M. D. PARTRIDGE (2004): "Do Urban Agglomeration Effects and Household Amenities Have a Skill Bias?," *Journal of Regional Science*, 44, 201–224.

ANDERSON, S. P., A. DE PALMA, AND J.-F. THISSE (1987): "The CES is a Discrete Choice Model?," *Economics Letters*, 24, 139–140.

ANDREWS, D. W. K. (1999): "Consistent Moment Selection Procedures for Generalized Method of Moments Estimation," *Econom*, 67(3), 543–564.

ASKER, J. (2004): "Measuring Cost Advantages from Exclusive Dealing: An Empirical Study of Beer Distribution," *Harvard mimeo*.

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

BEHRENS, K., AND F. ROBERT-NICOUD (2011): "Survival of the Fittest in Cities: Agglomeration, Selection, and Inequality," *CEPR Discussion Papers*, (7019).

BERNARD, A. B., J. B. JENSEN, S. J. REDDING, AND P. K. SCHOTT (2007): "Firms in International Trade," *The Journal of Economic Perspectives*, 21(3), pp. 105–130.

BILS, M., AND P. J. KLENOW (2001): "Quantifying Quality Growth," *The American Economic Review*, 91(4), 1006–1030.

BLACK, D., N. KOLESNIKOVA, AND L. J. TAYLOR (2009): "Earnings Functions When Wages and Prices Vary by Location," *Journal of Labor Economics*, 27, 21.

BRAKMAN, S., H. GARRETSEN, AND C. VAN MARREWIJK (2009): *The New Introduction to Geographical Economics*. Cambridge University Press.

BRODA, C., E. LIEBTAG, AND D. E. WEINSTEIN (2009): "The Role of Prices in Measuring the Poor's Living Standards," *Journal of Economic Perspectives*, 23(2), 77–97.

BRODA, C., AND J. ROMALIS (2009): "The Welfare Implications of Rising Price Disperison," .

BRODA, C., AND D. E. WEINSTEIN (2008): "Prices, poverty, and inequality: Why Americans are better off than you think," AEI Press.

——— (2010): "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, 100, 691.

COLEMAN, B. D., M. A. MARES, M. R. WILLIG, AND Y.-H. HSIEH (1982): "Randomness, Area, and Species Richness," *Ecology*, 63(4), 1121–1133.

COLWELL, R. K., AND J. A. CODDINGTON (1994): "Estimating Terrestrial Biodiversith through Extrapolation," *Philosophical Transactions: Biological Sciences*, 345(1311), 101–118.

COMBES, P.-P., G. DURANTON, AND L. GOBILLON (2008): "Spatial Wage Disparities: Sorting Matters!," *Journal of Urban Economics*, 63, 723–742.

COMBES, P.-P., G. DURANTON, L. GOBILLON, D. PUGA, AND S. ROUX (2009): "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," *CEPR Discussion Papers*, (7191).

COMBES, P.-P., T. MAYER, AND J.-F. THISSE (2008): *Economic Geography: the Integration of Regions and Nations.* Princeton University Press.

DAVIS, M. A., AND M. PALUMBO (2007): "The Price of Residential Land in Large US Cities," *Journal of Urban Economics*, 63(1), 352–384.

DE PALMA, A., AND K. KILANI (2007): "Invariance of Conditional Maximum Utility," *Journal of Economic Theory*, 132, 137–146.

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

DUMOND, J. M., B. T. HIRSCH, AND D. A. MACPHERSON (1999): "Wage Differentials Across Labor Markets and Workers: Does Cost of Living Matter?," *Economic Inquiry*, 37(4), 577–598.

FAJGELBAUM, P., G. M. GROSSMAN, AND E. HELPMAN (2011): "Income Distribution, Product Quality, and International Trade," *Journal of Political Economy*, 119(4), 721–765.

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

——— (2010): *Product Variety and the Gains from International Trade.* MIT Press.

FIELER, A. C. (2011): "Non-Homotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 79(4), 1069–1101.

FUJITA, M., AND T. MORI (2005): "Frontiers of the New Economic Geography," *Papers in Regional Science*, 84(3), 377–405.

GYOURKO, J., C. MAYER, AND T. SINAI (2006): "Superstar Cities," *NBER Working Paper 12355.*

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

HAUSMAN, J. A., G. LEONARD, AND J. D. ZONA (1994): "Competitive Analysis with Differentiated Products," *Annals D'Economie et Statistique*, 34, 159–180.

HELPMAN, E. (1998): *The Size of Regions*pp. 33–54. Cambridge University Press.

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

HUMMELS, D., AND P. J. KLENOW (2005): "The Variety and Quality of a Nation's Exports," *The American Economic Review*, 95(3), 704–723.

HUMMELS, D., AND V. LUGOVSKYY (2009): "International Pricing in a Generalized Model of Ideal Variety," *Journal of Money, Credit and Banking*, 41(1), 3–33.

JIMENEZ-VALVERDE, A., S. JIMENEZ MENDOZA, J. M. CANO, AND M. L. MUNGUIRA (2006): "Comparing relative model fit of several species-accumulation functions to local Papilionoidea and Hesperioidea butterfly inventories of Mediterranean habitats," *Biodiversity and Conservation*, 15, 163–176.

KHANDELWAL, A. (2010): "The Long and Short (of) Quality Ladders," *The Review*, 77(4), 1450–1476.

KRUGMAN, P. (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483–499.

LEE, S. (2010): "Ability Sorting and Consumer City," *Journal of Urban Economics*, 68, 20–33.

LI, N. (2011): "An Engel Curve for Variety," *Job Market Paper*.

MAO, C. X., R. K. COLWELL, AND J. CHANG (2004): "Interpolating, Extrapolating, and Comparing Incidence-Based Species Accumulation Curves," *Ecology*, 85(2717-2727), 433–441.

——— (2005): "Estimating the Species Accumulation Curve Using Mixtures," *Biometrics*, 61(2), 433–441.

MARKUSEN, J. R. (2010): "Putting Per-Capita Income Back into Trade Theory," *NBER Working Paper*, (15903).

MELITZ, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

MORETTI, E. (2008): "Real Wage Inequality," *NBER Working Paper 14370*.

MURPHY, K. M., AND R. H. TOPEL (1985): "Estimation and Inference in Two-Step Econometric Models," *Journal of Business & Economic Statistics*, 3(4), 370–379.

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

NEWEY, W. K., AND D. MCFADDEN (1994): "Large Sample Estimation and Hypothesis Testing," *Handbook of Econometrics*, IV, 2111–2245.

OF THE ROYAL SWEDISH ACADEMY OF SCIENCES, P. C. (2008): "Scientific background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2008," *http://nobelprize.org/nobel_prizes/economics/laureates/2008/ecoadv08.pdf*.

OTTAVIANO, G. I., T. TABUCHI, AND J.-F. THISSE (2002): "Agglomeration and Trade Revisited,," *International Economic Review*, 43(2), 409–435.

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

REDDING, S. J., AND D. M. STURM (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *The American Economic Review*, 98(5), 1766–1797.

ROBACK, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90(6), 1257–1278.

SATO, K. (1976): "The Ideal Log-Change Index Number," *The Review of Economics and Statistics*, 58(2), 223–228.

SIMONOVSKA, I. (2010): "Income Differences and Prices of Tradables," *UC Davis Working Paper*, No. 10-15.

SONG, I., AND P. K. CHINTAGUNTA (2007): "A Discrete-Continuous Model for Multicategory Purchase Behavior of Households," *Journal of Marketing Research*, 54, 595–612.

SUEDEKUM, J. (2006): "Agglomeration and Regional Costs of Living," *Journal of Regional Science*, 46(3), 529–543.

TABUCHI, T. (2001): "On Interregional Price Differentials," *The Japanese Economic Review*, 52(1), 104–115.

VARTIA, Y. (1976): "Ideal Log-Change Index Numbers," *Scandinavian Journal of Statistics*, 3(3), 121–126.

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

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

WUERGLER, T. (2010): "Income Distribution and Product Quality versus Variety," .