Supplemental Appendix E for

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

Jessie Handbury[1]

Contents

1 Non-Homotheticity Condition **1**

2 Connection to CES Utility Function **5**

3 Estimation Procedure **8**

4 Additional Results **13**

4.1 Full Distribution of Non-Homothetic Parameter Estimates . . . 13

4.2 Out-of-Sample Fit . . . 13

4.3 Non-Parametric Price Index Results . . . 15

5 Alternative Functional Form: CES Upper-Tier **17**

5.1 Model . . . 17

5.1.1 Individual Utility Maximization Problem . . . 17

5.1.2 Measuring Relative Utility Across Markets . . . 18

5.2 Parameter Estimation . . . 18

5.3 Results . . . 21

5.4 Appendices to CES Upper-Tier Analysis . . . 21

5.4.1 Derivation of Module-Level Expenditure Shares . . . 21

5.4.2 Between-Module Relative Market Expenditure Shares . . . 23

1 Non-Homotheticity Condition

Suppose 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:

$$f(U_{iG}(\mathbb{Q},Z),Z) \quad \text{subject to} \quad \sum_{m\in\mathbf{M}}\sum_{g\in\mathbf{G_m}} p_{mg}q_{mg} + Z \leq Y_i,\ q_{mg} \geq 0\ \forall\, mg \in \mathbf{G} \tag{E.1}$$

[1]handbury@wharton.upenn.edu

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

$$\max_{\mathbb{Q},Z} U_{iG}(\mathbb{Q}, Z) = \prod_{m\in M} \left[\sum_{g\in G_m} q_{mg} \exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img}) \right]^{\lambda_m}$$

$$\text{subject to} \quad \sum_{m\in \mathbf{M}} \sum_{g\in \mathbf{G_m}} p_{mg}q_{mg} \leq Y_i - Z, \; q_{mg} \geq 0 \,\forall\, mg \in \mathbf{G} \tag{E.2}$$

where $\gamma_m(Z) = (1 + \gamma_m \ln Z)$ and $\mu_m(Z) = \frac{1}{\alpha_m^0 + \alpha_m^1 \ln Z}$. Equations (7) and (E.18) define the optimal grocery bundle, $\mathbb{Q}^*(Z) = \left\{ \left\{ q_{mg}^*(Z) \right\}_{g\in\mathbf{G},} \right\}_{m\in\mathbf{M}}$ and can be summarized as follows:

$$q_{img}^*(Z) = \begin{cases} \frac{\lambda_m(Y_i - Z)}{p_{mg}} & \text{if } g = \arg\max_{g\in\mathbf{G_m}} \tilde{p}_{img} \\ 0 & \text{otherwise} \end{cases}$$

where

$$\tilde{p}_{img} = \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{ig})}{p_{mg}}$$

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

$$\begin{aligned}
\tilde{U}_{iG}(Z) &= U_{iG}(\mathbb{Q}^*(Z), Z) \\
&= \left\{ \sum_{m\in M} \left[\left((Y_i - Z) \frac{(\tilde{p}_{img})^\sigma}{P_i(Z)^{1-\sigma}} \right) \mathbb{I} \left[g = \arg\max_{g\in\mathbf{G_m}} \tilde{p}_{img} \right] \right]^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}} \\
&= \frac{Y_i - Z}{P_i(Z)^{1-\sigma}} \left\{ \sum_{m\in M} \left[\tilde{p}_{img}^\sigma \mathbb{I} \left[g = \arg\max_{g\in\mathbf{G_m}} \tilde{p}_{img} \right] \right]^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}} \\
&= \frac{Y_i - Z}{P_i(Z)^{1-\sigma}} \left\{ \sum_{m\in M} \left(\max_{g\in\mathbf{G_m}} \tilde{p}_{img} \right)^{\sigma-1} \right\}^{\frac{\sigma}{\sigma-1}} \\
&= \frac{Y_i - Z}{P_i(Z)}
\end{aligned} \tag{E.3}$$

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

$$\max_Z f(\tilde{U}_{iG}(Z), Z) \tag{E.4}$$

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

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

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

$$Y_i = Z - \frac{P_i(Z)}{P_i'(Z)} + \frac{f_2(\tilde{U}_{iG}(Z), Z)}{f_1(\tilde{U}_{iG}(Z), Z)} \frac{P_i(Z)^2}{P_i'(Z)}$$

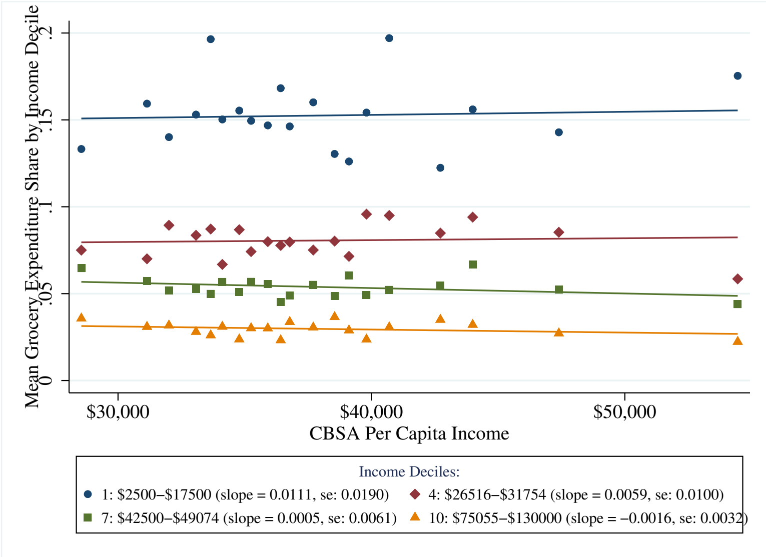
Taking the derivative of income with respect to non-grocery expenditure, Z, we can see that the non-grocery will be normal if the price vector and aggregate utility function are such that:

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

It is computationally infeasible to show that this condition holds generally (there will be a different price index $P_i(Z)$ for each of universe of potential price vectors), but I can show that it holds in the data by simply demonstrating that non-grocery expenditures are increasing in household income. I annualize the observed grocery expenditure for each household and measure annual non-grocery expenditures as the difference between the mid-point of each household's reported income category and the household's annual grocery expenditures. After controlling for household demographics with dummies for household size, marital status, education and age of the male and female heads of household, race, and Hispanic origin, the elasticity of observed non-grocery expenditures, Z_i, with respect to household income, Y_i, is 1.19 with a standard error of 0.003.

Figure E.1 demonstrates that households earning higher incomes spend a smaller share of their income on grocery products. Within income groups, however, the average grocery expenditure share does not vary much across cities and, in particular, Table E.I confirms that the average grocery share of an income group in a city does not vary systematically with city income.

Figure E.1: Income-Specific Grocery Expenditure Shares Across Markets



Note: Each point reflects the mean grocery expenditure share of households in each income decile that reside in households at each CBSAs at each vigntile of the CBSA per capita income distribution plotted against the mean CBSA per capital income of that vigntile. The household expenditure share is calculated as the annual reported expenditures on groceries (for households reporting trips in all 12 months of the year) divided by their reported income. For the purposes of visual clarity, only a representative sample of deciles are represented. The coefficient of variation of household grocery expenditure shares is 71 across all households in the sample, but drops to between 42 and 52 when you only consider households within each income decile. For the purposes of visual clarity, only a representative sample of deciles are represented.

Table E.I: Income-Specific Grocery Expenditure Shares Across Markets

	Dependent Variable: Mean Grocery Expenditure Share of Households in Income Decile									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Ln(CBSA PC Income)	0.011	-0.0034	-0.0045	0.0059	0.0048	-0.0051	0.00046	0.0075*	0.0060	-0.0016
	(0.019)	(0.012)	(0.013)	(0.010)	(0.0095)	(0.0072)	(0.0061)	(0.0042)	(0.0056)	(0.0032)
Observations	383	321	325	356	316	318	313	356	170	225

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. This table reports the correlation between the grocery expenditure share of Nielsen household panelists from each income decile and the per capita income of the CBSA where they reside. Observations are at the decile-by-CBSA level. The nth column reports regression for nth income decile.

2 Connection to CES Utility Function

In Section 4 of the paper, I model consumer demand assuming that a consumer i's utility from grocery consumption, conditional on their non-grocery expenditure Z, is a Cobb-Douglas aggregate over consumer-specific module-level utilities that are, in turn, additive in product-level log-logit utilities. This utility function is presented in equations (1), (2), and (3) and can be summarized as:

$$\begin{aligned} U_{iG}(\mathbb{Q}, Z) &= \prod_{m \in M} \left(u_{im}\left(\mathbb{Q}_m, Z\right)\right)^{\lambda_m} \\ &= \prod_{m \in M} \left(\sum_{g \in \mathbf{G_m}} u_{img}(\mathbb{Q}_m, Z) \right)^{\lambda_m} \\ &= \prod_{m \in M} \left(\sum_{g \in \mathbf{G_m}} q_{mg} \exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img}) \right)^{\lambda_m} \\ &= \prod_{m \in M} \left(\sum_{g \in \mathbf{G_m}} q_{mg} \exp(\gamma_m(Z)\beta_{mg} + \frac{\varepsilon_{img}}{\sigma_m(Z) - 1}) \right)^{\lambda_m} \end{aligned} \tag{E.5}$$

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

Consider the utility of the representative agent for consumers with non-grocery expenditure Z. This agent's utility function from grocery consumption is defined in equation (E.20) in Section 5.1 as follows:

$$U_G^{CES}(\mathbb{Q}, Z) = \prod_{m \in M} \left[\sum_{g \in \mathbf{G_m}} \left[q_{mg} \exp(\beta_{mg}\gamma_m(Z))\right]^{\frac{\sigma_m(Z)-1}{\sigma_m(Z)}} \right]^{\left(\frac{\sigma_m(Z)}{\sigma_m(Z)-1}\right)\lambda_m}, \tag{E.6}$$

where q_{mg}, β_{mg}, $\gamma_m(Z)$, $\sigma_m(Z)$, and λ_m take the same definitions as in equation (E.5) above.

Suppose that this representative consumer with the Cobb Douglas-nested CES utility function $U_G^{CES}(\mathbb{Q}, Z)$ defined in equation (E.6) faces the same prices $\mathbb{P}$ and has the same non-grocery expenditure Z as a group of "idiosyncratic" consumers with the Cobb Douglas-nested log-logit utility $U_{iG}(\mathbb{Q}, Z)$ defined in equation (E.5). A simple extension of Anderson et al. (1987) shows that the representative consumer and the group of "idiosyncratic" consumers will allocate expenditures across products within modules and across modules identically.

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

$$s_{mg|m}^{CES}(Z) = \left[\frac{\frac{p_{mg}}{\exp(\beta_{mg}\gamma_m(Z))}}{P_m^{CES}(Z, \mathbb{P}_m)} \right]^{1-\sigma_m(Z)}$$

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

$$\text{(E.7)} \qquad \ln s_{mg|m}^{CES}(Z) - \ln s_{m\bar{g}|m}^{CES}(Z) = (1 - \sigma_m(Z)) \left((\ln p_{mg} - \ln p_{m\bar{g}}) - (\beta_{mg} - \beta_{m\bar{g}})\gamma_m(Z) \right)$$

The expected relative module expenditure share of a group of "idiosyncratic" consumers with non-grocery expenditure Z facing the same prices p_{mg} and $p_{m\bar{g}}$ is derived in Appendix (C.2) as:
(E.8)

$$\mathbb{E}_\varepsilon \left[\ln(s_{img|m}(Z, \mathbb{P}_m)) - \ln(s_{im\bar{g}|m}(Z, \mathbb{P}_m)) \right] = (\sigma_m(Z) - 1) \left[(\beta_{mg} - \beta_{m\bar{g}})\gamma_m(Z) - (\ln p_{mg} - \ln p_{m\bar{g}}) \right]$$

where I have substituted $\sigma_m(Z)$ and $\gamma_m(Z)$ for their log-linear parametric forms $\left(1 + \alpha_m^0 + \alpha_m^1 \ln Z\right)$ and $(1 + \gamma_m \ln Z)$, respectively. We can multiply both terms of the right-hand side of (E.8) to show that it is equivalent to the right-hand side of equation (E.7):

$$\begin{aligned} \mathbb{E}_\varepsilon \left[\ln(s_{img|m}(Z, \mathbb{P}_m)) - \ln(s_{im\bar{g}|m}(Z, \mathbb{P}_m)) \right] &= (\sigma_m(Z) - 1) \left[(\beta_{mg} - \beta_{m\bar{g}})\gamma_m(Z) - (\ln p_{mg} - \ln p_{m\bar{g}}) \right] \\ &= (1 - \sigma_m(Z)) \left((\ln p_{mg} - \ln p_{m\bar{g}}) - (\beta_{mg} - \beta_{m\bar{g}})\gamma_m(Z) \right) \\ &= \ln s_{mg|m}^{CES}(Z) - \ln s_{m\bar{g}|m}^{CES}(Z) \end{aligned}$$

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

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

$$\text{(E.9)} \qquad \ln s_m^{CES}(Z) - \ln s_{\overline{m}}^{CES}(Z) = (1 - \sigma) \left(\ln \left(P_m^{CES}(Z, \mathbb{P}_m) \right) - \ln \left(P_{\bar{m}}^{CES}(Z, \mathbb{P}_{\bar{m}}) \right) \right)$$

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

$$\text{(E.10)} \qquad P_m^{CES}(Z, \mathbb{P}_m) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{p_{mg}}{\exp(\beta_{mg}\gamma_m(Z))} \right)^{(1-\sigma_m(Z))} \right]^{\frac{1}{(1-\sigma_m(Z))}}$$

The expected relative module expenditure share of a group of "idiosyncratic" consumers with non-grocery expenditure Z facing the same sets of prices $\mathbb{P}_m$ and $\mathbb{P}_{\bar{m}}$ faced by the representative consumer is derived in Appendix (5.4.2) as:

$$\text{(E.11)} \qquad \mathbb{E}_\varepsilon \left[\ln s_{im}(Z, \mathbb{P}) - \ln s_{i\bar{m}}(Z, \mathbb{P}) \right] \; = \; (\sigma - 1) \left[\ln V_m(Z, \mathbb{P}_m) - \ln V_{\bar{m}}(Z, \mathbb{P}_{\bar{\mathbf{m}}}) \right]$$

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

$$V_m(Z, \mathbb{P}_m) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\beta_{mg}\gamma_m(Z))}{p_{mg}} \right)^{(\sigma_m(Z)-1)} \right]^{\frac{1}{(\sigma_m(Z)-1)}} \tag{E.12}$$

To see that the right-hand sides of equations (E.9) and (E.11) are identical first note that we can re-write the equation (E.11) as

$$\begin{aligned} \mathbb{E}_\varepsilon \left[\ln s_{im}(Z, \mathbb{P}) - \ln s_{i\bar{m}}(Z, \mathbb{P}) \right] &= (1-\sigma) \left[-\ln V_m(Z, \mathbb{P}_m) + \ln V_{\bar{m}}(Z, \mathbb{P}_{\bar{\mathbf{m}}}) \right] \\ &= (1-\sigma) \left[\ln \left([V_m(Z, \mathbb{P}_m)]^{-1} \right) - \ln \left([V_{\bar{m}}(Z, \mathbb{P}_{\bar{\mathbf{m}}})]^{-1} \right) \right] \end{aligned}$$

In fact, the right-hand sides of equations (E.9) and (E.11) will be identical as long as the quality-adjusted price levels defined in equation (E.10) are equal to the inverse of the price-adjusted quality levels defined in equation (E.12), *i.e.*, $P_m^{CES}(Z, \mathbb{P}_m) = [V_m(Z, \mathbb{P}_m)]^{-1}$. We can see this is the case below:

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

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

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

3 Estimation Procedure

In this appendix I describe the details involved in the estimation and statistical inference of the lower-level parameters, $\theta_1 = \{\alpha_m^0, \alpha_m^1, \gamma_m, \{\beta_{mg} - \beta_{m\bar{g}_m}\}\}_{m=1,...,M}$. This set of demand parameters is partitioned into M sets of lower-level module-specific parameters, θ_{1m} for each module m, that are identified using module-specific sub-samples of the data. The upper-level parameters – Cobb-Douglas module expenditure weights, $\theta_2 = \{\lambda_m\}_{m=1,...,M}$ – is calibrated to the module-level sales shares in the estimation sample.

I obtain $\hat{\theta}_1$ using a two-stage GMM procedure based on the following exogeneity restriction:

$$\mathbb{E}[g(\mathbf{X};\theta_1)] = 0 \tag{E.13}$$

where $g(\mathbf{X};\theta_1)$=$[g^1(\mathbf{X};\theta), g^2(\mathbf{X};\theta), g^3(\mathbf{X};\theta)]$ consists of three vectors of module-specific moments, $g^k(\mathbf{X};\theta) = [g^k(\mathbf{X}_1;\theta_1), ..., g^k(\mathbf{X}_M;\theta_M)]$.

The first vector of moments is calculated using market-level data. They are defined as:

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

where n is the number of product-market observations; $\xi_{mgt}(\mathbf{X}_m;\theta^{NL}_{1m})$ are transient market-specific product taste shocks defined below; and $\mathbf{Z}^1_{mgt}$ is a vector of L^1_m pre-determined variables including product fixed effects and price instruments. The tilde denotes that a variable has been differenced from the respective value for the base product in each module, $\bar{g}_m$, e.g., $\tilde{\xi}_{mgst}(\mathbf{X}_m;\theta_{1m}) = \xi_{mgst}(\mathbf{X}_m;\theta_{1m}) - \xi_{m\bar{g}_m st}(\mathbf{X}_m;\theta_{1m})$.

The second and third vectors of moments are designed to employ the Nielsen data on household-level product choices. The second set of moments equalizes the predicted uncentered covariance between product quality and household non-grocery expenditure for Nielsen HMS sample households. The sample analog of this covariance is:

$$\bar{g}^2(\mathbf{X_m};\theta_{1m}) = \frac{1}{N_m}\sum_{mg} g^2_{mg}(\mathbf{X_m};\theta_{1m}) = \frac{1}{N_m}\sum_{mg} N_{mg}\beta_{mg}\left\{\frac{1}{N_{mg}}\sum_{i_{mg}=1}^{n_{mg}} Z_{i_{mg}} - E[Z|y=mg,\theta]\right\}$$

where i_{mg} denotes one of the N_{mg} units of product g in module m that is purchased in the Nielsen HMS sample; i denotes one of the N_m units of any product in module m that is purchased in the Nielsen HMS sample; and Z_i denotes the non-grocery expenditure of the Nielsen HMS panelist purchasing unit i. Similarly, the third set of moments equalizes the predicted uncentered covariance between unit price paid and household non-grocery expenditure. The sample analog of this covariance is:

$$\hat{\bar{g}}^3(\mathbf{X}_m;\theta_{1m}) = \frac{1}{N_m}\sum_i \left(Z_i - \bar{Z}\right)\sum_t \left((\tilde{p}_{imt} - E[\tilde{p}_{imt}|\theta_{1m}]) - \frac{1}{N_m}\sum_i\sum_t (\tilde{p}_{imt} - E[\tilde{p}_{imt}|\theta_{1m}])\right)$$

The sample analogs of the three moment conditions defined above are:

$$
\begin{aligned}
\hat{\tilde{g}}^1(\mathbf{X}_m;\theta_{1m}) &= \frac{1}{n}\sum_{mg,t}\hat{\tilde{\xi}}_{mgt}(\mathbf{X}_m;\theta_{1m})\tilde{\mathbf{Z}}_{mgt} \\
\hat{\tilde{g}}^2(\mathbf{X}_m;\theta_{1m}) &= \frac{1}{N_m}\sum_{mg}N_{mg}\left\{\beta_{mg}\left[\frac{1}{N_{mg}}\sum_{i_{mg}=1}^{N_{mg}}Z_{i_{mg}}-\frac{1}{N_m}\sum_{i=1}^{N_m}Z_iP_{mg}(Z_i,\mathbb{P}_t,\theta_{1m},\hat{\beta}_t)\right]\right\}^2 \\
\hat{\tilde{g}}^3(\mathbf{X}_m;\theta_{1m}) &= \frac{1}{N_m}\sum_i\left(Z_i-\bar{Z}\right)\sum_t\left((\tilde{p}_{imt}-E[\tilde{p}_{imt}|\theta_{1m}])-\frac{1}{N_m}\sum_i\sum_t\left(\tilde{p}_{imt}-E[\tilde{p}_{imt}|\theta_{1m}]\right)\right)
\end{aligned}
$$

where $\bar{Z}=\frac{1}{N_m}\sum_i\bar{Z}_i$ is the unit-weighted mean non-grocery expenditure of sample households; $\tilde{p}_{imt}=(p_{imgt}-\bar{p}_{mt})$ is the relative unit value paid by a household i in module m in market t, where $\bar{p}_{mt}=\sum_{g\in\mathbf{G}_{mt}}w_{mgt}p_{mgt}$ and $w_{mgt}=s_{mg}/\sum_{g\in\mathbf{G}_{mt}}s_{mg}$, and $E[\tilde{p}_{imt}|\theta_{1m}]$ is the predicted relative unit value paid by household i in module m in market t defined as:[2]

$$
E[\tilde{p}_{imt}|\theta_{1m}] = \sum_{g\in\mathbf{G}_{mst}}\tilde{p}_{mgt}P_{mg}(Z_i,\mathbb{P}_t,\theta_{1m},\hat{\beta}_t)
$$

To obtain estimates for the quality parameters $\tilde{\beta}_{mg}(\theta_{1m}^{NL})$ that enter the micro moments, I first follow Berry et al. (1995) inverting simulated market shares to obtain the vector product- and market-specific taste parameters $\tilde{\beta}_{mgt}(\theta_{1m}^{NL})$ that rationalizes the observed product shares in each market conditional on a given set of non-linear parameter vector $\theta_{1m}^{NL}=\left\{\alpha_m^0,\alpha_m^1,\gamma_m\right\}$. Details on the simulation and inversion procedure are provided below.[3] I project the estimated taste parameters, $\hat{\tilde{\xi}}_{mgt}(\theta_{1m}^{NL})$, on brand as well as market dummies to control for market-level variation in the quality of the products included in the base good. The coefficients on the brand dummies are used as estimates for the product-specific quality parameters, $\tilde{\beta}_{mg}(\theta_{1m}^{NL})$, employed in the quality micro moment. The residuals from these regressions provide estimates for the transitory shocks, $\xi_{mgt}(\theta_{1m}^{NL})$, which are in turn used to calculate the macro (store-level) moment conditions.

The fact that all three sets of moments depend only on module-specific data, $\mathbf{X}_m$, and parameters, θ_{1m}, enables me to partition E.13 into module-specific auxiliary moments:

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

This partition allows me to estimate the K_{1m} parameters, $\theta_{1m}=\left\{\alpha_m^0,\alpha_m^1,\gamma_m,\{\tilde{\beta}_{mg}\}_{g\in\mathbf{G_m},g\neq\bar{g}_m}\right\}$, for each module m in separate but parallel minimization procedures. Consistent estimates of the elas-

[2] I can only calculate the probability of purchase, $P_{mg}(Z_i,\mathbb{P}_t,\theta_{1m},\hat{\beta}_t)$, employed in the calculation of the micro moments ($\hat{\tilde{g}}^2(\mathbf{X}_m;\theta_{1m})$ and $\hat{\tilde{g}}^3(\mathbf{X}_m;\theta_{1m})$), when I observe the full choice set available to the Nielsen household panelist i; that is, the set of products and prices available to the customer in the store and time period that they are observed to make their purchase ($\mathbb{P}_t$). I observe these choice sets for the stores and time periods in the Nielsen RMS data, so calculate the micro moments using household transactions in these stores and time periods alone.

[3] I also attempted estimating these taste shocks using a fourth set of moments equalizing the predicted expenditure shares of a simulated set of customers at each store in each time period with the observed sales shares for the respective stores and time periods following Dubé et al. (2012)'s implementation of Berry et al. (1995). I ran into difficulties getting this model to converge across many modules, however, given the non-linearity of the problem.

ticity parameters, $\theta_{1m}^{NL} = \left\{\alpha_m^0, \alpha_m^1, \gamma_m\right\}$, are obtained by minimizing module-specific GMM objective functions as follows:

$$\hat{\theta}_{1m}^{NL} = \arg\min_{\theta_{1m}^{NL}} \hat{\bar{g}}(\mathbf{X}_m;\theta_{1m})'\hat{\mathbf{W}}_{1m}\hat{\bar{g}}(\mathbf{X}_m;\theta_{1m})$$

where $\hat{\bar{g}}(\mathbf{X}_m;\theta_{1m})$ is the sample analog of the $L_m^1 + 1 \geq K_{1m}$ moments, $\bar{g}(\mathbf{X}_m;\theta_{1m})$ and $\hat{\mathbf{W}}_{1m}$ is the efficient weighting matrix.

The weighting matrix, $\hat{\mathbf{W}}_{1m}^1$, is block-diagonal since the three moments are calculated using different datasets:

$$\hat{\mathbf{W}}_{1m}^1 = \begin{bmatrix} \hat{W}_{1m}^{11}(\mathbf{X}_m;\tilde{\theta}_{1m}) & 0 & 0 \\ 0 & \hat{W}_{1m}^{12}(\mathbf{X}_m;\tilde{\theta}_{1m}) & 0 \\ 0 & 0 & \hat{W}_{1m}^{13}(\mathbf{X}_m;\tilde{\theta}_{1m}) \end{bmatrix}^{-1}$$

for

$$\begin{aligned} \hat{W}_{1m}^{11}(\mathbf{X}_m;\tilde{\theta}_{1m}) &= \frac{1}{n}\sum_{mg,t} \hat{g}_{mgt}^1(\mathbf{X}_m;\tilde{\theta}_{1m})\hat{g}_{mgt}^1(\mathbf{X}_m;\tilde{\theta}_{1m})' \\ \hat{W}_{1m}^{12}(\mathbf{X}_m;\tilde{\theta}_{1m}) &= \frac{1}{N_m}\sum_{mg} \hat{g}_{mg}^2(\mathbf{X}_m;\tilde{\theta}_{1m})\hat{g}_{mg}^2(\mathbf{X}_m;\tilde{\theta}_{1m})' \\ \hat{W}_{1m}^{13}(\mathbf{X}_m;\tilde{\theta}_{1m}) &= \frac{1}{N_m}\sum_{mg} \hat{g}_{mg}^3(\mathbf{X}_m;\tilde{\theta}_{1m})\hat{g}_{mg}^3(\mathbf{X}_m;\tilde{\theta}_{1m})' \end{aligned}$$

Each of these components is calculated using consistent first-stage estimates of θ_{1m}^{NL}:

$$\tilde{\theta}_{1m}^{NL} = \arg\min_{\theta_{1m}^{NL}} \hat{\bar{g}}(\mathbf{X}_m;\theta_{1m})'\mathbf{W}_{1m}\hat{\bar{g}}(\mathbf{X}_m;\theta_{1m})$$

for

$$\mathbf{W}_{1m} = \begin{bmatrix} \left[\frac{1}{n}\sum_{mg,t}\sum_{g\in\mathbf{G}_{mt}} \tilde{Z}_{mgt}^1\left(\tilde{Z}_{mgt}^1\right)'\right]^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

After estimating the non-linear parameters, $\hat{\theta}_{1m}^{NL}$, I project the product-store-time specific taste shocks implied by these parameters, $\tilde{\beta}_{mgt}(\hat{\theta}_{1m}^{NL})$, onto brand dummies in order to extract estimates of the product quality parameters, $\{\tilde{\beta}_{mg}\}_{g\in\mathbf{G_m},g\neq\bar{g}_m}$.

Assuming that the random components of the M module-specific auxiliary models are independent, the variance-covariance matrix of $\hat{\theta}_1$, Ω_1, can be written as:

$$\Omega_{\theta_1} = \begin{bmatrix} \Omega_{\theta_{11}} & & & & 0 \\ & \ddots & & & \\ & & \Omega_{\theta_{1m}} & & \\ & & & \ddots & \\ 0 & & & & \Omega_{\theta_{1M}} \end{bmatrix}$$

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

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

where $\hat{F}_{\theta_{1m}} = \begin{bmatrix} \hat{F}^1_{\theta_{1m}} & \hat{F}^2_{\theta_{1m}} \end{bmatrix}'$ for

$$\hat{F}^1_{\theta_{1m}} = \frac{1}{n} \sum_{mg,t} \nabla_{\theta_{1m}} \hat{g}^1_{mgt}(\mathbf{X}_m; \hat{\theta}_{1m})$$

and

$$\hat{F}^2_{\theta_{1m}} = \frac{1}{N_m} \sum_{mg} \nabla_{\theta_{1m}} \hat{g}^2_{mg}(\mathbf{X}_m; \hat{\theta}_{1m})$$

and

$$\hat{V}_{ff} = \begin{bmatrix} \frac{1}{n} \sum_{mg,t} \hat{g}^1_{mgt}(\mathbf{X}_m; \hat{\theta}_{1m}) \hat{g}^1_{mgt}(\mathbf{X}_m; \hat{\theta}_{1m})' & 0 \\ 0 & \frac{1}{N_m} \sum_{mg} \hat{g}^2_{mg}(\mathbf{X}_m; \hat{\theta}_{1m}) \hat{g}^2_{mg}(\mathbf{X}_m; \hat{\theta}_{1m})' \end{bmatrix}$$

Inversion Algorithm In order to evaluate the objective function at a given parameter vector θ^{NL}_{1m}, it is necessary to invert the following system of non-linear equations:

(E.14) $$\beta_{mgt}(\theta_{1m}) \rightarrow \ln s_{mgt}(\beta_t; \theta^{NL}_{1m}) = \ln \hat{s}_{mgt}$$

where $s_{mgt}(\beta_t; \theta^{NL}_{1m})$ is the model predicted market share of product g in market t, $\theta^{NL}_{1m} = \{\alpha^0_m, \alpha^1_m, \gamma_m\}$ is the subset of elasticity parameters that must be estimated using non-linear moments, and $\hat{s}_{mgt}$ is the observed share. For each guess of θ^{NL}_{1m}, I calculate the model predicted market share as the average probability of purchase predicted for a quadrature of K points from the market-specific income distribution (recall that income is used to proxy for non-grocery expenditure Z_i) each with income Y_k and weight w_k:

(E.15) $$s_{mgt}(\beta_t; \theta^{NL}_{1m}) = \sum_{k=1}^{K} w_k P_{mg}(Y_k, \mathbb{P}, \theta_m)$$

It is well known that this inversion does not work for products with small sales shares (see, e.g., Gandhi et al. (2019)). I therefore group all of the products that fall into the left tail of the average sales distribution an outside product. This grouping could impact my estimates in three ways. First, Gandhi et al. (2019) have demonstrated that ignoring the low end of the sales distribution in this manner yields a downward bias on price elasticity estimates. Second, variation in the quality of the outside goods sold in different stores could bias my average product quality estimates as discussed under identification in Section 5.3.1. Finally, I will not estimate product quality parameters for products that always appear in the low end of the sales distribution and, therefore, am unable to include them in the market price indexes. To test the impact of these biases on my results, I study how the estimated price elasticities and product quality gradients vary depending on the share of products that are grouped into this outside product, varying this set between 40, 60, and 80 percent of products in each store-week (reflecting 6, 15, and 33 percent of aggregate product sales, respectively) in the robustness exercises presented in Section 6.4.1.

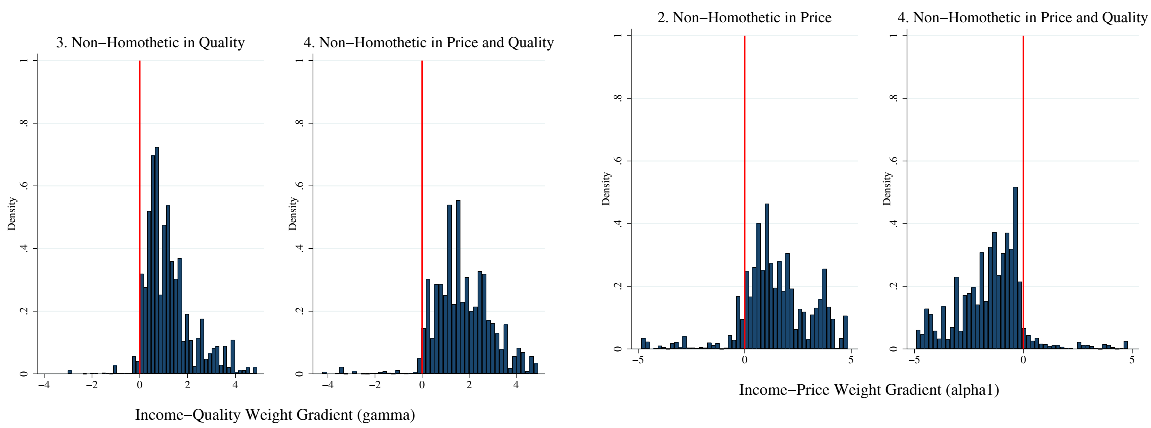
Starting Values I estimate a linear approximation of the store-level market share equation to obtain starting values for the non-linear parameters, $\theta_{1m}^{NL} = \left\{\alpha_m^0, \alpha_m^1, \gamma_m\right\}$. When the optimization routine returns estimates within 0.03 log units of the bounds for these non-linear estimates –$\alpha_m^0 \in (0.05, 30)$, $\alpha_m^1 \in (-5, 5)$, and $\gamma_m \in (-5, 5)$ – or otherwise fails, I instead conduct a grid search. Specifically, I run the optimization routine using a range of starting values for the mean price elasticity, $\alpha_m^{0,start}$ between 1 and 4, keeping the starting values for the non-homotheticity parameters of $\gamma_m^{start} = 1.5$ and $\alpha_m^{1,start} = 2$ (or zero, in the constrained model). If this yields multiple sets of interior estimates, I select the estimates minimize the objective function.

4 Additional Results

4.1 Full Distribution of Non-Homothetic Parameter Estimates

Figure E.2: Distribution of Parameter Estimates Across Modules

a. Gamma (γ_m) b. Alpha 1 (α_m^1)



Notes: The left-hand plot above depicts the distribution of the γ_m estimates, for the model allowing for non-homotheticity in the demand for quality alone (i.e., restricting that α_m^1=0) on the left and for the model allowing non-homotheticity in both the demand for quality and price sensitivity (i.e., allowing both γ_m and α_m^1 to be non-zero) on the right. The right-hand plot above depicts the distribution of the α_m^1 estimates, for the model allowing for non-homotheticity in price sensitivity alone (i.e., restricting that γ_m =0) on the left and for the model allowing non-homotheticity in both the demand for quality and price sensitivity (i.e., allowing both γ_m and α_m^1 to be non-zero) on the right. Attention is limited to modules for which the estimation procedure converged at interior estimates for all relevant parameters.

4.2 Out-of-Sample Fit

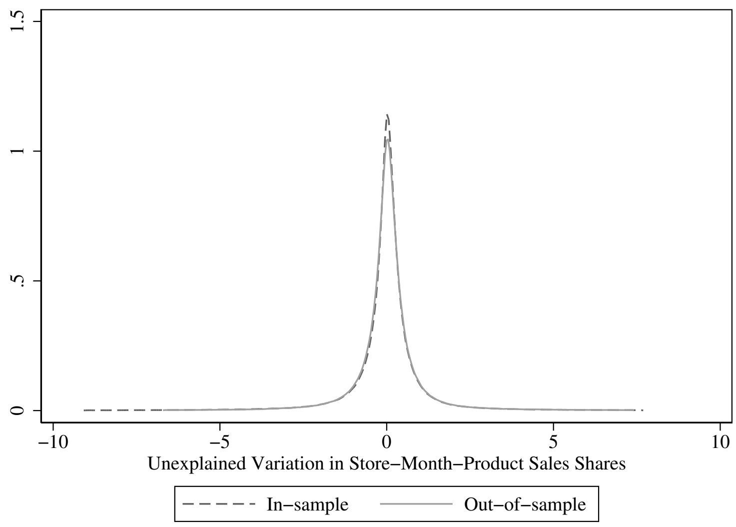
The model is currently estimated using data describing sales in a random sample of 1000 CBSA-month markets for each product module. This leaves plenty of data to study the out-of-sample fit. The analysis below studies the out-of-sample fit for the baseline model used for the price index analysis (i.e., the model that allows non-homotheticity in the demand for quality, but not price sensitivity).

Figure E.3 compares the distribution of the unexplained component of store-month sales, which take the structural interpretation of transient taste shocks, in the estimation sample with that in a secondary sample of 1000 CBSA-month markets for each product module. The two distributions–truncated at the 1st and 99th percentiles–are very similar to one another.

This fit is summarized in the J-statistics of the macro moments.[4] Figure E.4 compares the J-statistics calculated using the model estimates for α_m^0 and γ_m in the secondary sample to the J-statistics for the estimation sample. The average fit is, as expected, worse out-of-sample, but, barring some outliers, the fit of the macro moments is highly correlated across modules between the estimation and secondary samples.

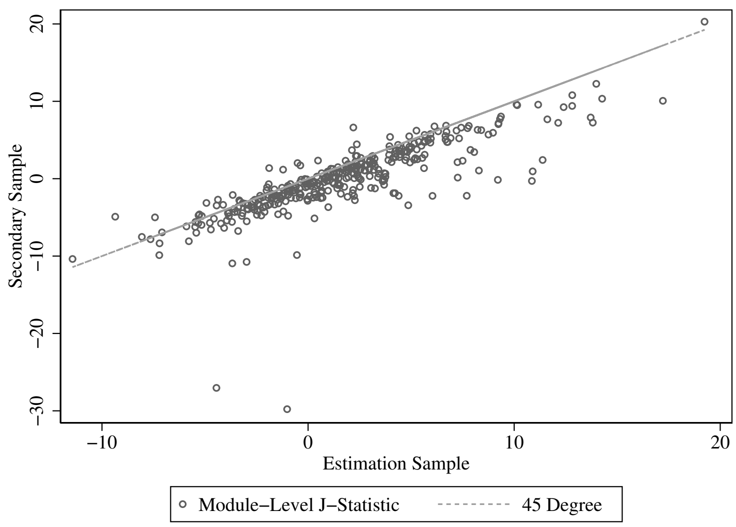
[4]The CBSA-month sampling procedure prioritizes CBSA-months where HMS households are observed to make product purchases, so there is not a secondary sample of household purchases with which I can calculate out-of-sample micro moments.

Figure E.3: Transient Taste Shocks ($\xi_{mgt} - \beta_{mg}$) Predicted In-Sample and Out-of-Sample



Notes: This plot shows the distribution of transient CBSA-month tastes for products, estimated using sales in the base sample of 1000 CBSA-month markets (in-sample) and then calculated using the same non-linear parameter estimates in a hold-back sample of 1000 different CBSA-month markets (out-of-sample). This out-of-sample check is for the baseline model that allows for non-homotheticity in the demand for quality but not in price sensitivity (i.e., restricting that α_m^1=0).

Figure E.4: J-Statistics for CBSA-Level Moments In-Sample and Out-of-Sample



Notes: This plot compares the fit of the CBSA-level moments estimated using sales in the base sample of 1000 CBSA-month markets (the "estimation" sample) and then calculated using the same non-linear parameter values but for a hold back sample of 1000 different CBSA-month markets (the "secondary" sample) across different modules. The fit of these moments in each sample is summarized with a module-level J statistic calculated with the weighting matrix and CBSA-level moment conditions described above in Appendix Section 3. This out-of-sample check is for the baseline model that allows for non-homotheticity in the demand for quality but not in price sensitivity (i.e., restricting that α_m^1=0).

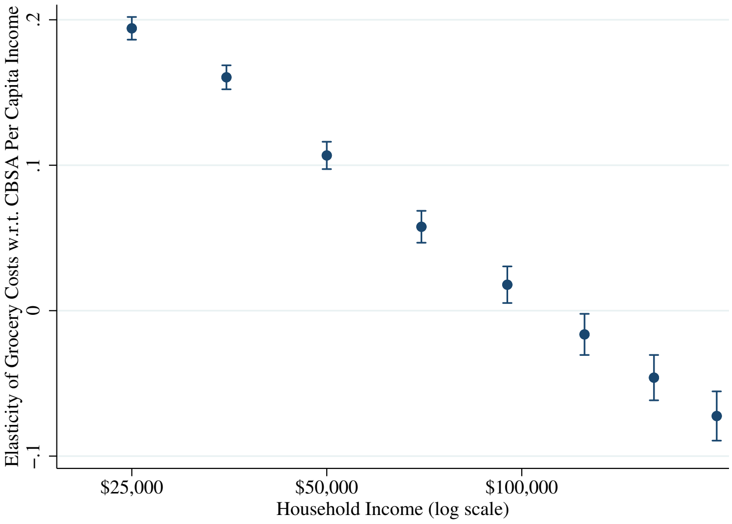
4.3 Non-Parametric Price Index Results

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

$$\ln \hat{P}(\mathbb{P}_c, y_k) = \delta_k + \beta_1 y_c + \beta_{2k} y_c + \epsilon_{kc}, \tag{E.16}$$

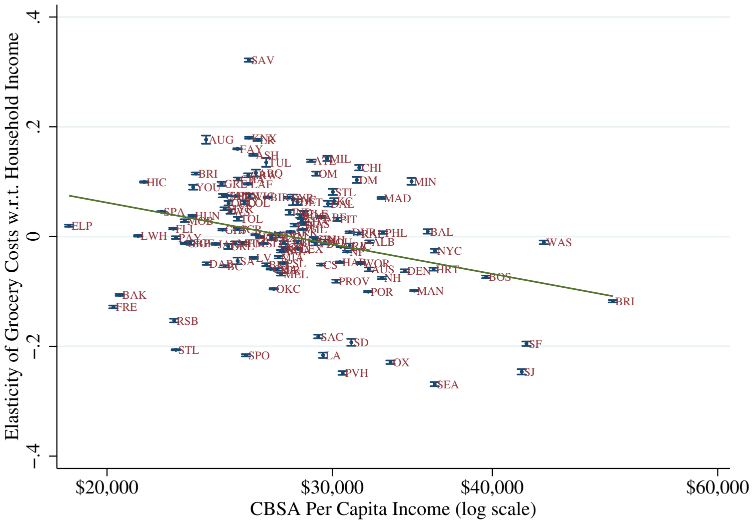
I estimate this regression separately for each set of 100 bootstrapped samples of 50 random stores from each CBSA. Figure E.5 plots the mean of the resulting β_{2k} elasticity parameter estimates against log household income, y_k. These results indicate that there is indeed a linear relationship between this elasticity and household income. Figure E.6 further shows the log linear relationship between the semi-elasticity of price indexes with respect to market income and CBSA income; i.e., β_{2c} in $\ln \hat{P}(\mathbb{P}_c, y_k) = \delta_c + \beta_1 y_k + \beta_{2c} y_k + \epsilon_{kc}$.

Figure E.5: Variation Across Bootstrap Samples in the Elasticity of Grocery Price Index with respect to CBSA Income for Households at Different Size-Adjusted Income Levels



Notes: This plot shows the elasticity of income- and CBSA-specific price indexes with respect to CBSA per capita income for households at compares the different income levels. The point shows the mean elasticity estimated across 100 bootstrap iterations of price index calculations (each drawing a random sample of 50 stores in each CBSA) and the bands show the 95 percent confidence intervals around this mean. The price indexes are calculated using the baseline model that allows for non-homotheticity in the demand for quality but not in price sensitivity (i.e., restricting that α_m^1=0).

Figure E.6: Variation Across Bootstrap Samples in the Elasticity of Grocery Price Index with respect to Household Income for CBSAs with Different Per Capita Income



Notes: This plot shows the elasticity of household income- and CBSA-specific price indexes with respect to household income in CBSAs with different per capita incomes. The point shows the mean elasticity estimated across 100 bootstrap iterations of price index calculations (each drawing a random sample of 50 stores in each CBSA) and the bands show the 95 percent confidence intervals around this mean. The price indexes are calculated using the baseline model that allows for non-homotheticity in the demand for quality but not in price sensitivity (i.e., restricting that α_m^1=0). The marker labels for each CBSA are acronyms linked to the full CBSA name in Appendix A.4.

5 Alternative Functional Form: CES Upper-Tier

In the main text of the paper, I assume that substitution between product modules is governed by Cobb-Douglas utility. In this appendix, I present the model, estimation procedure, and results under the alternative assumption of CES utility. The results are similar to the baseline because the estimated elasticity of substitution between modules is close to one.

5.1 Model

Under CES demand, a consumer i's utility from grocery consumption, conditional on their non-grocery expenditure Z, is a CES aggregate over consumer-specific module-level utilities:

$$U_{iG}(\mathbb{Q}, Z) = \left\{ \sum_{m \in M} u_{im} \left(\mathbb{Q}_m, Z \right)^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}} \tag{E.17}$$

where $\sigma > 1$ is the elasticity of substitution between modules and module-level utility is as defined in the main text (equations (2) and (3)).

5.1.1 Individual Utility Maximization Problem

Consumers then solve for their optimal grocery consumption bundle for a given non-grocery expenditure level Z by maximizing grocery utility subject to budget and non-negativity constraints (equation (6)). The solution to this problem is a vector of optimal product selections (one for each module), $\mathbf{g}_\mathbf{i}^*(Z) = (g_{i1}^*(Z), ..., g_{iM}^*(Z))$, and module-level expenditures, $\mathbf{w}_\mathbf{i}^*(Z) = (w_{i1}^*(Z), ..., w_{iM}^*(Z))$. The optimal product selections are invariant to the upper-tier utility assumption, so defined as in equation (7) in the main text:

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

The optimal module-level expenditures under the CES assumption are derived in appendix (5.4.1) below to be:

$$w_{im}^*(Z) \quad = \quad (Y_i - Z) \frac{\left(\max\limits_{g \in \mathbf{G_m}} \left(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img} \right) / p_{mg} \right)^{\sigma-1}}{P(\mathbb{P}, Z, \varepsilon_\mathbf{i})^{1-\sigma}} \tag{E.18}$$

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

$$P(\mathbb{P}, Z, \varepsilon_\mathbf{i}) = \left[\sum_{m \in \mathbf{M}} \left(\max_{g \in \mathbf{G_m}} \left(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img} \right) / p_{mg} \right)^{\sigma-1} \right]^{\frac{1}{1-\sigma}} \tag{E.19}$$

5.1.2 Measuring Relative Utility Across Markets

I measure relative grocery costs across cities using the price index faced by a representative consumer. The representative consumer's utility from consuming a grocery bundle $\mathbb{Q}$ is a nested-CES function conditional on their non-grocery expenditure Z defined as:

$$U_G^{CES}(\mathbb{Q}, Z) = \left\{ \sum_{m \in M} \left[\sum_{g \in \mathbf{G_m}} \left[q_{mg} \exp(\beta_{mg} \gamma_m(Z)) \right]^{\frac{\sigma_m(Z)-1}{\sigma_m(Z)}} \right]^{\left(\frac{\sigma_m(Z)}{\sigma_m(Z)-1}\right)\left(\frac{\sigma-1}{\sigma}\right)} \right\}^{\frac{\sigma}{\sigma-1}}, \tag{E.20}$$

In appendix 2 below , I show that this income-specific, nested, asymmetric CES utility function yields identical within-grocery budget shares as the CES-nested log-logit utility function that I estimate.

The indirect utility of this representative consumer from income Y_i and prices and products $\mathbb{P}_t$, $V^{CES}(\mathbb{P}_t, Y_i)$, can be expressed as the ratio of the consumer's grocery expenditure to a price index that summarizes the consumer's marginal utility from expenditure given the prices and products available in the market:

$$V^{CES}(\mathbb{P}_t, Y_i, Z_{it}) = \frac{(Y_i - Z_{it})}{P^{CES}(\mathbb{P}_t, Z_{it})}, \tag{E.21}$$

where

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

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

5.2 Parameter Estimation

The routine for estimating the parameters that govern demand allocations across products within modules (θ_1) are unchanged from that presented in section 5.3 of main text. The parameters that goven cross-module expenditure allocations with the CES upper-tier are the cross-module substitution parameter, σ, and the quality of the base product in each module, $\beta_{m\bar{g}_m}$, for all modules $m \in \mathbf{M}$, except for the base module $\bar{m}$.[5] I denote this set of parameters by θ_2:

$$\theta_2 = \left\{ \sigma, \{\beta_{m\bar{g}_m}\}_{m \in \mathbf{M}, m \neq \bar{m}} \right\}$$

To estimate these parameters, I use a single set of moments that fit the predicted store-level module sales shares observed in the Nielsen RMS data to those predicted by the model.

The expected log expenditure share in module m relative to $\bar{m}$ for a group of households with the same non-grocery expenditure, Z_i, facing a common vector of grocery prices, $\mathbb{P}$, is derived below in Appendix 5.4.1. Adjusting this expression to reflect time-varying market-specific pricing and promotion

[5] I normalize the fixed quality of the base product in the base module (butter), $\beta_{\bar{m}\bar{g}_{\bar{m}}}$, to equal zero.

activity yields:

$$\mathbb{E}_{\varepsilon}\left[\ln s_{imt} - \ln s_{i\bar{m}t}\right] = (\sigma - 1) \ln \tilde{V}_m(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t}) \tag{E.22}$$

where $\tilde{V}_{mt}(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t}) = V_{mt}(Z_i, \mathbb{P}_{mt}) / V_{\bar{m}t}(Z_i, \mathbb{P}_{\bar{m}t})$. $V_{mt}(Z_i, \mathbb{P}_{mt})$ is a CES-style index over price-adjusted product qualities:

$$V_m(Z_i, \mathbb{P}_{mt}) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\gamma_{im}\beta_{mgt})}{p_{mgt}}\right)^{-\alpha_{im}}\right]^{\frac{1}{-\alpha_{im}}} \tag{E.23}$$

Note that the inclusive value is a function of the parameters estimated in both the first and second stage, i.e., θ_1 and θ_2. To see this recall that $\alpha_{im} = \left(\alpha_m^0 + \alpha_m^1 \ln Z_i\right)$ and $\gamma_{im} = (1 + \gamma_m \ln Z_i)$ and each market-specific product quality shock, β_{mgt}, is the the sum of $(\beta_{mgt} - \beta_{m\bar{g}_m t})$, estimated in stage 1, and an unknown base product quality shock, $\beta_{m\bar{g}_m t}$. We can express the inclusive value function as the product of the base product quality parameter, $\beta_{m\bar{g}_m t}$, to be estimated in the second stage and an inclusive value function calculated using only elements of θ_{1m} estimated in the first stage:

$$V_m(Z_i, \mathbb{P}_{mt}) = \exp(\gamma_{im}\beta_{m\bar{g}_m t}) V_{1m}(Z_i, \mathbb{P}_{mt})$$

where

$$V_{1m}(Z_i, \mathbb{P}_{mt}) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\gamma_{im}\tilde{\beta}_{mgt})}{p_{mgt}}\right)^{-\alpha_{im}}\right]^{\frac{1}{-\alpha_{im}}} \tag{E.24}$$

and $\tilde{\beta}_{mgt} = \beta_{mgt} - \beta_{m\bar{g}_m t}$. Under the normalization that $\beta_{\bar{m}\bar{g}_{\bar{m}} t} = 0$ for all t, and using the decomposition of the inclusive value function above, we can now rewrite equation (E.22) as:

$$\mathbb{E}_{\varepsilon}\left[\ln s_{imt} - \ln s_{i\bar{m}t}\right] = (\sigma - 1)\left(\gamma_{im}\beta_{m\bar{g}_m t} + \ln \tilde{V}_{1mt}(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t})\right) \tag{E.25}$$

where $\ln \tilde{V}_{1mt}(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t}) = \ln V_{1m}(Z_i, \mathbb{P}_{mt}) - \ln V_{1\bar{m}}(Z_i, \mathbb{P}_{\bar{m}t})$.

The predicted log expenditure share of module m relative to module $\bar{m}$ in market t is obtained by aggregating i-specific expected relative shares over the units purchased by customers at each non-grocery expenditure level:

$$\mathbb{E}_z\left[\mathbb{E}_{\varepsilon}\left[\ln s_{imt} - \ln s_{i\bar{m}t}\right]\right] = \beta_{m\bar{g}_m t}(\sigma - 1)\overline{\gamma}_{mt} + (\sigma - 1)\overline{\tilde{v}}_{mt} \tag{E.26}$$

where $\overline{\gamma}_{mt} = \int \gamma_{im} dF(Z|t)$ and $\overline{\tilde{v}}_{mt} = \int \ln \tilde{V}_{1mt}(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t}) dF(Z|t)$ can be calculated using price data and parameter estimates for θ_1 obtained in stage 1 above.

The moment equation is then defined as:

$$\bar{h}(\theta_2) = \frac{1}{n}\sum_{m,t} h_{mt}(\theta_2) = \frac{1}{n}\sum_{m,t} u_{mt}(\mathbf{X}; \hat{\theta}_1, \theta_2)\mathbf{W}_{mt}$$

where n is the number of (market-module) observations; $\mathbf{W}_{mt}$ includes the average market-level quality coefficient $\overline{\gamma}_{mt}$ interacted with module fixed effects and an instrument for the average relative inclusive value for the module, $\overline{\overline{v}}_{mt}$, described below; and u_{mt} denotes the difference between the observed log relative module shares between modules m and $\bar{m}$ in market t and their predicted values, i.e.,

$$u_{mt}(\mathbf{X};\hat{\theta}_1,\theta_2) = \ln\left(s_{mt}/s_{\bar{m}t}\right) - \beta_{m\bar{g}_m}\left(\sigma-1\right)\overline{\gamma}_{mt}(\hat{\theta}_1) - \left(\sigma-1\right)\overline{\overline{v}}_{mt}(\hat{\theta}_1) \tag{E.27}$$

Identification of σ and $\beta_{m\bar{g}_m}$ relies on the assumption that the errors in the model predicted shares (u_{mt}) are orthogonal from $\mathbf{W}_{mt}$. The u_{mt} errors can be broken into two components, $u_{mt} = u^1_{mt} + u^2_{mt}$. The first, $u^1_{mt} = (\sigma-1)\left(\beta_{m\bar{g}_m}\left(\overline{\gamma}_{mt} - \overline{\gamma}_{mt}(\hat{\theta}_1)\right) + \overline{\overline{v}}_{mt} - \overline{\overline{v}}_{mt}(\hat{\theta}_1)\right)$, reflect errors in the first stage estimates, while the second, $u^2_{mt} = \xi_{m\bar{g}_m t}\left(\sigma-1\right)\overline{\gamma}_{mt}$ for $\xi_{m\bar{g}_m t} = \beta_{m\bar{g}_m t} - \beta_{m\bar{g}_m}$, reflect the transitory components of the product-market taste shocks that are not estimated directly. To deal with the endogeneity of prices with respect to these transitory taste shocks, I instrument for the average inclusive value, $\overline{v}_{mt}$, using a data analog calculated with the same contemporaneous chain-specific national cost shock instruments that are used in the module-level estimation in place of market-specific price data.

The σ substitution elasticity parameter is identified by the extent to which relative module shares react to national chain-specific cost shocks for each module. Recall that the relative inclusive value, $\overline{\overline{v}}_{mt}$, is scaled up or down by the quality of the base product, $\bar{g}_m$, in a module m relative to the quality of the base product, $\bar{g}_{\bar{m}}$, in the base module $\bar{m}$, butter (a product type sold in most stores), which is normalized to equal zero. Any difference between the expenditure share of module m relative to butter and what would be expected given the relative inclusive value of the two modules and the σ estimate will identify the quality of the base product in the module, $\beta_{m\bar{g}_m}$, scaled by the market average taste for quality, $\overline{\gamma}_{mst}$. Together with the relative product quality estimates from the first stage of estimation, $\beta_{mg} - \beta_{m\bar{g}_m}$, the base product quality estimates define the quality of each product in the dataset relative to the quality of the base product in the base module.

The upper-level estimation yields between-module elasticity σ estimates reported in Table E.II. As expected, products in different modules are less substitutable than products in the same module, with between-module substitution elasticities close to one.

Table E.II: Upper-Level Substitution Elasticity Estimates

Model Name	σ
Homothetic	1.007
	[0.137]
Non-Homothetic in Price	1.019
	[0.162]
Non-Homothetic in Quality	1.002
	[0.004]
Non-Homothetic in Quality and Price	1.001
	[0.000]

Note: This table shows the estimates for the elasticity of substitution between modules.

5.3 Results

Table E.III compares the main result of the paper using the baseline indexes that assume Cobb-Douglas upper-tier (in columns [1] and [2]) with these results using indexes assuming a CES upper-tier (in columns [3] and [4]). The cross-elasticity of grocery costs with respected to city and household income estimated without controls is higher with CES demand (-0.26 in column [3] vs. -0.20 in column [1]) but the difference is not statistically significant and the two coefficients converge once population controls are added in columns [2] and [4]. This is not surprising, given how close the estimated elasticity of substitution parameters in Table E.II are to one.

Table E.III: City-Income Specific Price Index Regressions using CES Upper-Tier

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

Across-Module Aggregation:	Cobb-Douglas (Baseline)		CES	
	[1]	[2]	[3]	[4]
Ln(Per Capita Income$_c$) (β_1)	-0.068	-0.042	-0.10	-0.031
	(0.088)	(0.10)	(0.13)	(0.15)
Ln(Per Capita Income$_c$)*	-0.18***	-0.15***	-0.26***	-0.19**
Demeaned Ln(HH Income$_k$) (β_2)	(0.038)	(0.039)	(0.061)	(0.061)
Ln(Population$_c$) (β_3)		-0.0095		-0.026
		(0.018)		(0.026)
Ln(Population$_c$)*		-0.011		-0.026**
Demeaned Ln(HH Income$_k$) (β_4)		(0.0072)		(0.0094)
Income Group k*Bootstrap Sample FEs	Yes	Yes	Yes	Yes
Number of CBSAs (c)	125	125	125	125
Observations	100,000	100,000	100,000	100,000
adj. within R^2	0.02	0.02	0.02	0.03

Notes: *** p< 0.01, ** p<0.05, * p<0.10; standard errors, clustered by bootstrap sample and CBSA, are in parentheses.This table presents results from regressions of household income- and CBSA-specific grocery price indexes against CBSA characteristics alone and interacted with demeaned log household income. The price indexes correspond to the model that allows for non-homotheticity in the demand for quality but not in price sensitivity (i.e., restricting that α_m^1=0) and measure how households at eight different income levels between \$25,000 and \$200,000 value the products and prices represented in each of 100 bootstrap samples of 50 stores in each of 125 CBSAs with 50 or more participating retailers. The price indexes studied in columns [1] and [2] assume Cobb-Douglas upper-tier demand system, as presented in the main text. The price indexes studied in columns [3] and [4] assume CES upper-tier demand, as described in the appendix above.

5.4 Appendices to CES Upper-Tier Analysis

5.4.1 Derivation of Module-Level Expenditure Shares

Consumer i, spending Z on non-grocery items, chooses how to allocate expenditures between modules by selecting $w_1, ..., w_M$ to maximize

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

subject to

$$\sum_{m\in\mathbf{M}} w_m \leq Y_i - Z$$

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

$$\max_{g\in\mathbf{G_m}} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})}{p_{mg}} = \frac{1}{A_{im}} \tag{E.28}$$

The within-module allocation decision now simplifies to:

$$\mathbf{w_i^*}(Z) = (w_{i1}^*(Z), ..., w_{iM}^*(Z)) = \underset{\sum_{m\in\mathbf{M}} w_m \leq Y_i - Z}{\arg\max} \left\{\sum_{m\in M}\left[\frac{w_m}{A_{im}}\right]^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{\sigma}{\sigma-1}} \tag{E.29}$$

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:

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

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':

$$\begin{aligned}
\left\{\sum_{m\in M}\left[\frac{w_m}{A_{im}}\right]^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{1}{1-\sigma}} \frac{1}{A_{im'}}\left[\frac{w_{m'}}{A_{im'}}\right]^{-\frac{1}{\sigma}} &= \left\{\sum_{m\in M}\left[\frac{w_m}{A_{im}}\right]^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{1}{1-\sigma}} \frac{1}{A_{im}}\left[\frac{w_m}{A_{im}}\right]^{-\frac{1}{\sigma}} \\
\frac{1}{A_{im'}}\left[\frac{w_{m'}}{A_{im'}}\right]^{-\frac{1}{\sigma}} &= \frac{1}{A_{im}}\left[\frac{w_m}{A_{im}}\right]^{-\frac{1}{\sigma}} \\
w_{m'} &= w_m\left[\frac{A_{im'}}{A_{im}}\right]^{1-\sigma}
\end{aligned}$$

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

$$
\begin{aligned}
Y_i - Z &= \sum_{m' \in \mathbf{M}} w_{m'} \\
Y_i - Z &= \frac{w_m}{A_{im}^{1-\sigma}} \sum_{m' \in \mathbf{M}} [A_{im'}]^{1-\sigma} \\
w_m &= \frac{A_{im}^{1-\sigma}}{\sum\limits_{m' \in \mathbf{M}} [A_{im'}]^{1-\sigma}} (Y_i - Z)
\end{aligned}
$$

The solution to problem (E.29) is, therefore,

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

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

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

Substituting from equation (E.28) for A_{img} yields consumer i's optimal module expenditure vector, $\mathbf{w_i^*}(Z)$, as a function of total grocery expenditures, prices, and model parameters:

$$\mathbf{w_i^*}(\mathbf{Z}) = (w_{i1}^*(Z), ..., w_{iM}^*(Z)) \;\; \text{where} \;\; w_{im}^* = (Y_i - Z) \frac{\left[\max\limits_{g \in \mathbf{G_m}} \dfrac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})}{p_{mg}} \right]^{\sigma-1}}{P_i(Z)^{1-\sigma}}$$

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

5.4.2 Between-Module Relative Market Expenditure Shares

I now want to generate a estimatiing equation that can be used to identify σ and $\{\beta_{\bar{g}_m}\}_{g \in \mathbf{G_m}}$ using data on module-level income-specific market shares. The optimal cross-module expenditure allocation for consumer i conditional on this consumer's idiosyncratic utility draws for each product in each module is characterized by the following equations:

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

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

where $\tilde{p}_{img} = \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})}{p_{mg}}$. Dividing through by total grocery expenditure, $(Y_i - Z)$, I generate consumer i's optimal module m expenditure share, conditional on their non-grocery expenditure Z and the vector of prices they face, $\mathbb{P}$:

$$s_{im}(Z, \mathbb{P}) = \frac{w^*_{im}(Z)}{Y_i - Z} = \frac{\left[\max\limits_{g \in \mathbf{G_m}} \tilde{p}_{img}\right]^{\sigma - 1}}{P_i^{1-\sigma}}$$

When deriving the within-module relative market share, I take the expectation of the consumer's expected product expenditure share over the idiosyncratic errors, $\mathbb{E}_\varepsilon[s_{img|m}(Z, \mathbb{P}_m)]$, to derive an expression for the market share of each product. I then divide these market shares by the market share of a module specific base product and take 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.[6]

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

$$\frac{s_{im}(Z, \mathbb{P})}{s_{i\bar{m}}(Z, \mathbb{P})} = \frac{\left[\max\limits_{g \in \mathbf{G_m}} \tilde{p}_{img}\right]^{\sigma - 1}}{\left[\max\limits_{g \in \mathbf{G_{\bar{m}}}} \tilde{p}_{i\bar{m}g}\right]^{\sigma - 1}}$$

Since P_i does not vary across modules for a given consumer i, it drops out of the relative module

[6]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:

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

I derive the expression for the Z-specific market share of product g, $s_{mg|m}(Z, \mathbb{P}_m) = \mathbb{E}_\varepsilon[s_{img|m}(Z, \mathbb{P}_m)]$, before taking logs and differencing to generate the estimation equation, as it demonstrates the relationship between the term on the left-hand side of this equation, $\ln(s_{mg|m}(Z, \mathbb{P}_m)) - \ln(s_{m\bar{g}|m}(Z, \mathbb{P}_m))$, and its value in the data: the difference between the log of the expenditure consumers spending Z on non-grocery items in a given market on product g relative to the log of their expenditure on the base product $\bar{g}$ or, more succinctly, the log difference between the Z-specific market shares on products g and $\bar{g}$.

expenditure share expression. I take the log of this relative share expression to linearize the equation:

$$\ln s_{im}(Z,\mathbb{P}) - \ln s_{i\bar{m}}(Z,\mathbb{P}) = (\sigma-1)\ln\left(\max_{g\in\mathbf{G_m}} \tilde{p}_{img}\right) - (\sigma-1)\ln\left(\max_{g\in\mathbf{G_{\bar{m}}}} \tilde{p}_{i\bar{m}g}\right),$$

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:

(E.30)

$$\mathbb{E}_\varepsilon\left[\ln s_{im}(Z,\mathbb{P}) - \ln s_{i\bar{m}}(Z,\mathbb{P})\right] = (\sigma-1)\left\{\mathbb{E}_\varepsilon\left[\ln\left(\max_{g\in\mathbf{G_m}} \tilde{p}_{img}\right)\right] - \mathbb{E}_\varepsilon\left[\ln\left(\max_{g\in\mathbf{G_{\bar{m}}}} \tilde{p}_{i\bar{m}g}\right)\right]\right\}$$

Consider the two expectation terms in equation (E.30). Both take the same form, and thus I only solve for the first expectation:

$$\mathbb{E}_\varepsilon\left[\ln\left(\max_{g\in\mathbf{G_m}} \tilde{p}_{img}\right)\right] \tag{E.31}$$

The expectation term defined in equation (E.31) 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:

$$\begin{aligned}
\mathbb{E}_\varepsilon\left[\ln\left(\max_{g\in\mathbf{G_m}} \tilde{p}_{img}\right)\right] &= \mathbb{E}_\varepsilon\left[\ln\left(\max_{g\in\mathbf{G_m}} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)\right] \\
&= \mathbb{E}_\varepsilon\left[\max_{g\in\mathbf{G_m}} \ln\left(\frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)\right] \\
&= \mathbb{E}_\varepsilon\left[\max_{g\in\mathbf{G_m}} \gamma_m(Z)\beta_{mg} - \ln p_{mg} + \mu_m(Z)\varepsilon_{img}\right] \\
&= \mu_m(Z)\mathbb{E}_\varepsilon\left[\max_{g\in\mathbf{G_m}} (\gamma_m(Z)\beta_{mg} - \ln p_{mg})/\mu_m(Z) + \varepsilon_{img}\right] && \text{(E.32)}
\end{aligned}$$

De Palma and Kilani (2007) show that, for an additive random utility model with $u_i = \nu_i + \varepsilon_i$, $i = 1,\ldots,n$ 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 (E.32) takes the form $\mathbb{E}_\varepsilon[\max_g \nu_{img} + \varepsilon_{img}]$, with $\nu_{img} = (\gamma_m(Z)\beta_{mg} - \ln p_{mg})/\mu_m(Z)$, 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:

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

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

$$\begin{aligned}
\mathbb{E}_{\varepsilon}\left[\max_{g\in\mathbf{G_m}} v_g+\varepsilon_g\right] &= \int_{-\infty}^{\infty} z\left(\sum_{g=1}^{G_m}\exp(v_g-z)\right)\exp\left(\sum_{g=1}^{G_m}-\exp(v_g-z)\right)dz\\
&= \int_{-\infty}^{\infty} z\exp\left(\sum_{g=1}^{G_m}-\exp(v_g-z)\right)\left(\sum_{g=1}^{G_m}\exp(v_g-z)\right)dz\\
&= \int_{\infty}^{0}(V-\ln x)\exp(-x)\,x(-1/x)dx\\
&= \int_{0}^{\infty}(V-\ln x)\exp(-x)\,dx\\
&= V
\end{aligned}$$

Since we have defined $\nu_{img}=(\gamma_m(Z)\beta_{mg}-\ln p_{mg})/\mu_m(Z)$ and $V=\ln\left[\sum_{g=1}^{G_m}\exp(v_g)\right]$, we can use the above result to solve for the expectation in equation (E.31):

$$\begin{aligned}
\mathbb{E}_{\varepsilon}\left[\ln\left(\max_{g\in\mathbf{G_m}}\tilde{p}_{img}\right)\right] &= \mu_m(Z)\ln\left[\sum_{g\in\mathbf{G_m}}\exp((\gamma_m(Z)\beta_{mg}-\ln p_{mg})/\mu_m(Z))\right]\\
&= \mu_m(Z)\ln\left[\sum_{g\in\mathbf{G_m}}\left(\frac{\exp(\gamma_m(Z)\beta_{mg})}{p_{mg}}\right)^{\frac{1}{\mu_m(Z)}}\right]\\
&= \ln\left[\sum_{g\in\mathbf{G_m}}\left(\frac{\exp(\gamma_m(Z)\beta_{mg})}{p_{mg}}\right)^{\frac{1}{\mu_m(Z)}}\right]^{\mu_m(Z)} \quad \text{(E.33)}
\end{aligned}$$

Plugging this result back into equation (E.30) yields the expected relative module expenditure share for consumer i in terms of product prices and model parameters:

$$
\begin{aligned}
\mathbb{E}_\varepsilon & \left[\ln s_{im}(Z,\mathbb{P}) - \ln s_{i\bar{m}}(Z,\mathbb{P})\right] \\
&= (\sigma-1)\mathbb{E}_\varepsilon\left[\ln\left(\max_{g\in\mathbf{G_m}} \frac{\exp(\gamma_m(Z)\beta_{mg} + \mu_m(Z)\varepsilon_{img})}{p_{mg}}\right)\right] \\
&\quad - (\sigma-1)\mathbb{E}_\varepsilon\left[\ln\left(\max_{g\in\mathbf{G_{\bar{m}}}} \frac{\exp(\gamma_{\bar{m}}(Z)\beta_{\bar{m}g} + \mu_{\bar{m}}(Z)\varepsilon_{i\bar{m}g})}{p_{\bar{m}g}}\right)\right] \\
&= (\sigma-1)\ln\left[\sum_{g\in\mathbf{G_m}}\left(\frac{\exp(\gamma_m(Z)\beta_{mg})}{p_{mg}}\right)^{\frac{1}{\mu_m(Z)}}\right]^{\mu_m(Z)} \\
&\quad - (\sigma-1)\ln\left[\sum_{g\in\mathbf{G_{\bar{m}}}}\left(\frac{\exp(\gamma_{\bar{m}}(Z)\beta_{\bar{m}g})}{p_{\bar{m}g}}\right)^{\frac{1}{\mu_{\bar{m}}(Z)}}\right]^{\mu_{\bar{m}}(Z)}
\end{aligned}
$$

This function only varies by consumer through their non-grocery expenditure. All consumers with the same non-grocery expenditure and facing the same prices, $\mathbb{P}$, will have the same expected relative module expenditure share:

$$
\text{(E.34)} \qquad \mathbb{E}_\varepsilon\left[\ln s_{im}(Z,\mathbb{P}) - \ln s_{i\bar{m}}(Z,\mathbb{P})\right] = -(\sigma-1)\left[\ln V_m(Z,\mathbb{P}_m) - \ln V_{\bar{m}}(Z,\mathbb{P}_{\bar{\mathbf{m}}})\right]
$$

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

$$
\text{(E.35)} \qquad V_m(Z,\mathbb{P}_m) = \left[\sum_{g\in\mathbf{G_m}}\left(\frac{\exp(\beta_{mg}(1+\gamma_m\ln Z))}{p_{mg}}\right)^{(1-\sigma)}\right]^{\frac{1}{1-\sigma}}
$$

where I have substituted in the parametrizations for $\gamma_m(Z) = (1+\gamma_m \ln Z)$ and $\mu_m(Z) = 1/\left(\alpha_m^0 + \alpha_m^1 \ln Z\right)$. Equations (E.34) and (E.35) 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 $g \in G_{\bar{m}}$.

Extracting Second Stage Estimates θ_2 From the Inclusive Value Function The expected log expenditure share in module m relative to $\bar{m}$ for a group of households with the same non-grocery expenditure, Z_i, facing a common vector of grocery prices, $\mathbb{P}$, is defined above in Equations (E.34) and (E.35). Adjusting these expressions to reflect time-varying CBSA-specific pricing and promotion activity yields:

$$
\text{(E.36)} \qquad \mathbb{E}_\varepsilon\left[\ln s_{imt} - \ln s_{i\bar{m}t}\right] = (\sigma-1)\ln\tilde{V}_m(Z_i,\mathbb{P}_{mt},\mathbb{P}_{\bar{m}t})
$$

where $\tilde{V}_m(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t}) = V_m(Z_i, \mathbb{P}_{mt})/V_{\bar{m}}(Z_i, \mathbb{P}_{\bar{m}t})$. $V_m(Z_i, \mathbb{P}_{mt})$ is a CES-style index over price-adjusted product qualities:

$$V_m(Z_i, \mathbb{P}_{mt}) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\gamma_{im}\beta_{mgt})}{p_{mgt}}\right)^{-\alpha_{im}}\right]^{\frac{1}{-\alpha_{im}}} \tag{E.37}$$

for $\alpha_{im} = \left(\alpha_m^0 + \alpha_m^1 \ln Z_i\right)$ and $\gamma_{im} = (1 + \gamma_m \ln Z_i)$. Note that the inclusive value is a function of the parameters estimated in both the first and second stage, i.e., θ_1 and θ_2. Specifically, each market-specific product quality shock, β_{mgt}, is the the sum of $(\beta_{mgt} - \beta_{m\bar{g}_m t})$, estimated in stage 1, and an unknown base product quality shock, $\beta_{m\bar{g}_m t}$. We can express the inclusive value function as the product of the base product quality parameter, $\beta_{m\bar{g}_m t}$, to be estimated in the second stage and an inclusive value function calculated using only elements of θ_{1m} estimated in the first stage:

$$V_m(Z_i, \mathbb{P}_{mt}) = \exp(\gamma_{im}\beta_{m\bar{g}_m t}) V_{1m}(Z_i, \mathbb{P}_{mt})$$

where

$$V_{1m}(Z_i, \mathbb{P}_{mt}) = \left[\sum_{g \in \mathbf{G_m}} \left(\frac{\exp(\gamma_{im}\tilde{\beta}_{mgt})}{p_{mgt}}\right)^{-\alpha_{im}}\right]^{\frac{1}{-\alpha_{im}}} \tag{E.38}$$

and $\tilde{\beta}_{mgt} = \beta_{mgt} - \beta_{m\bar{g}_m t}$. Under the normalization that $\beta_{\bar{m}\bar{g}_{\bar{m}} t} = 0$ for all t, and using the decomposition of the inclusive value function above, we can now rewrite equation (E.36) as:

$$\mathbb{E}_\varepsilon \left[\ln s_{imt} - \ln s_{i\bar{m}t}\right] = (\sigma - 1)\left(\gamma_{im}\beta_{m\bar{g}_m t} + \ln \tilde{V}_{1mt}(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t})\right) \tag{E.39}$$

where $\ln \tilde{V}_{1mt}(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t}) = \ln V_{1m}(Z_i, \mathbb{P}_{mt}) - \ln V_{1\bar{m}}(Z_i, \mathbb{P}_{\bar{m}t})$.

The predicted log expenditure share of module m relative to module $\bar{m}$ in market t is obtained by aggregating i-specific expected relative shares over the units purchased by customers at each non-grocery expenditure level:

$$\mathbb{E}_z \left[\mathbb{E}_\varepsilon \left[\ln s_{imt} - \ln s_{i\bar{m}t}\right]\right] = \int (\sigma - 1)\left(\gamma_{im}\beta_{m\bar{g}_m t} + \ln \tilde{V}_{1mt}(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t})\right) dF(Z|t) \tag{E.40}$$

where $F(Z|t)$ is the distribution of non-grocery expenditures over the households shopping in market t.

Notice that this function is linear in the unobserved base product quality for module m, $\beta_{m\bar{g}_m t}$, and the relative inclusive value function, so we can derive the following linear estimating equation:

$$\mathbb{E}_z \left[\mathbb{E}_\varepsilon \left[\ln s_{imt} - \ln s_{i\bar{m}t}\right]\right] = \beta_{m\bar{g}_m t}\,(\sigma - 1)\,\overline{\gamma}_{mt} + (\sigma - 1)\,\overline{\tilde{v}}_{mt} \tag{E.41}$$

where $\overline{\gamma}_{mt} = \int \gamma_{im} dF(Z|t)$ and $\overline{\tilde{v}}_{mt} = \int \ln \tilde{V}_{1mt}(Z_i, \mathbb{P}_{mt}, \mathbb{P}_{\bar{m}t}) dF(Z|t)$ can be calculated using price data, estimates of the market-level income distributions, and stage 1 parameter estimates.

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

$$G = \mathbb{E}[h(\mathbf{X};\theta_1,\theta_2)] = 0 \tag{E.42}$$

where $h(\mathbf{X};\theta_1,\theta_2) = \mathbf{Z}_2(\mathbf{X})\cdot u(\mathbf{X};\theta_1,\theta_2)$. $\mathbf{Z}_2(\mathbf{X})$ is a set of L_2 instruments ($L_2 \geq K_2$) and $u(\mathbf{X};\theta_1,\theta_2)$ is the error in the relative across-module expenditure share equation derived above.

Specifically, for module m and store s in time t this error is derived above in equation (E.41) as:

$$u_{mt}(\mathbf{X};\theta_1,\theta_2) \quad = \quad \ln\left(s_{mt}/s_{\bar{m}t}\right) - \beta_{m\bar{g}_m}\left(\sigma-1\right)\overline{\gamma}_{mt}(\hat{\theta}_1) - \left(\sigma-1\right)\overline{\tilde{v}}_{mt}(\hat{\theta}_1)$$

where s_{mt} and $s_{\bar{m}t}$ are data on the respective sales shares of module m and $\bar{m}$ in market t; each $\bar{x}_{mt} = \int x_{imt} dF(Z|t)$ is calculated by integrating x_{imt} over the same local income distribution employed in the first-stage of estimation described above, for $\gamma_{im} = (1 + \gamma_m \ln Z_i)$ and $\tilde{v}_{mt} = \ln V_{1m}(Z_i, \mathbb{P}_{mt}, \theta_{1m}) - \ln V_{1\bar{m}}(Z_i, \mathbb{P}_{\bar{m}t}, \theta_{1m})$ where

$$V_{1m}(Z_i, \mathbb{P}_{mt}, \theta_{1m}) = \left[\sum_{g\in\mathbf{G_m}} \left(\frac{\exp(\gamma_{im}\tilde{\beta}_{mgt})}{p_{mgt}}\right)^{-(\alpha_m^0+\alpha_m^1 \ln Z_i)}\right]^{\frac{1}{-(\alpha_m^0+\alpha_m^1 \ln Z_i)}}$$

is the inclusive value for a household with non-grocery expenditure Z_i in module m in market t calculated using first-stage parameter estimates, $\hat{\theta}_1$.

$\mathbf{Z}_2(\mathbf{X})$ is a vector of pre-determined variables including module fixed effects interacted with the market average quality weight, $\overline{\gamma}_{mt}$, and an instrument for the average relative inclusive value, $\overline{\tilde{v}}_{mt}(\hat{\theta}_1)$, faced by the store's customers. This instrument is identical to the data analog of $\overline{\tilde{v}}_{mt}(\hat{\theta}_1)$ but calculated using the same contemporaneous chain-specific national cost shock instruments that are used in the module-level estimation in place of market-specific price data.

The upper-level parameters are estimated using two-step GMM:

$$\hat{\theta}_2 = \arg\min_{\theta_2} \hat{h}(\mathbf{X};\hat{\theta}_1,\theta_2)'\hat{\mathbf{W}}_\mathbf{2}\hat{h}(\mathbf{X};\hat{\theta}_1,\theta_2)$$

where $\hat{\mathbf{W}}_2 = \left[\frac{1}{\sum_t N_t}\sum_t\sum_{m\in\mathbf{M}_t} h_{mt}(\mathbf{X};\hat{\theta}_1,\tilde{\theta}_2)h_{mt}(\mathbf{X};\hat{\theta}_1,\tilde{\theta}_2)'\right]^{-1}$ is the optimal weighting matrix, for $\tilde{\theta}_2$ the consistent first-stage estimates of θ_2 that minimize a GMM objective function as follows:

$$\tilde{\theta}_2 = \arg\min_{\theta_2} \hat{h}(\mathbf{X};\hat{\theta}_1,\theta_2)'\tilde{\mathbf{W}}_\mathbf{2}\hat{h}(\mathbf{X};\hat{\theta}_1,\theta_2)$$

where $\tilde{\mathbf{W}}_2 = \left[\frac{1}{\sum_t N_t} \sum_t \sum_{m \in \mathbf{M}_t} \mathbf{Z}_{2mt} \mathbf{Z}'_{2mt} \right]^{-1}$.

References

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

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

de Palma, A. and K. Kilani (2007). Invariance of Conditional Maximum Utility. *Journal of Economic Theory 132*, 137–146.

Dubé, J.-P., J. T. Fox, and C.-L. Su (2012). Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. *Econometrica 80*(5), 2231–2267.

Gandhi, A., Z. Lu, and X. Shi (2019). Estimating Demand for Differentiated Products with Zeroes in Market Share Data.