Online Appendix: The Demand Model and Model Identification

The Demand Model

Random Coefficients

For our demand random coefficient logit model, we follow the specifications in the framework of Berry, Levinsohn and Pakes (1995; see also Nevo 2000), BLP hereafter. The term ε_{ijt} in Equation 1 follows an i.i.d. type I extreme value distribution. We assume that the random coefficients β_i , α_i ,and γ_i follow multivariate normal distributions

$$\text{(A1)} \qquad \begin{pmatrix} \beta_i \\ \alpha_i \\ \gamma_i \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \\ \gamma \end{pmatrix} + \sum v_i \quad v_i \sim P_v(v)$$

where β, α, and γ measure the mean preference that is common among all consumers; v_i represents the unobserved variation of preferences assumed to have a standard multivariate normal distribution $P_v(v)$; and Σ is a [(K + 1) × (1 + K)] scaling matrix of the random coefficients that need to be estimated.

Let θ_1 = (β,α,γ) be the vector containing the linear parameters and θ_2 = Σ be the nonlinear parameters. Combining Equations 1 and A1, we have the indirect utility expressed as in Equation A2, where δ_{jt} refers to the mean utility that is common to all consumers. The term μ_{ijt} + ε_{ijt} depicts a mean-zero heteroskedastic deviation from the mean utility.

$$\text{(A2)} \qquad U_{ijt} = \delta_{jt}\big(X_{jt}, P_{jt}, S_{jt}, \xi_{jt}; \theta_1\big) + \mu_{ijt}\big(X_{jt}, P_{jt}, S_{jt}, v_i; \theta_2\big) + \varepsilon_{ijt}$$

$$\delta_{jt} = X_{jt}\beta + \alpha P_{jt} + \gamma S_{jt} + \xi_{ji} \qquad \mu_{ijt} = \big(X_{jt}, P_{jt}, S_{jt}\big) * \sum v_i$$

Consumer Choice and Market Share

We normalize the indirect utility of nonpurchase decision to zero, $U_{i0t} = \varepsilon_{i0t}$. Because we define ε_{ijt} as having a type I extreme value distribution, we have a closed-form solution for the probability that a consumer would purchase brand *j* in market *t*:

$$\text{(A3)} \qquad Pr_{ijt}\left(\delta_{jt}|\theta_2\right) = \frac{\exp\left(\delta_{jt}+\mu_{jit}\right)}{1+\sum_{m=1}^{J}\exp(\delta_{mt}+\mu_{mit})}$$

Aggregating over consumers, we can specify the market share of brand *j* in market *t* as follows:

$$\text{(A4)} \qquad S_{jt}\left(\delta_{jt}|\theta_2\right) = \int \frac{\exp\left(\delta_{jt}+\mu_{jt}\right)}{1+\sum_{m=1}^{J}\exp(\delta_{mt}+\mu_{mt})} dP_v(v) \ = \int Pr_{ijt}\, d\Psi(\alpha_i, \gamma_i\, |\theta_2)$$

where Ψ is the joint distribution of consumer characteristics, and θ_2 is a vector of parameters for this joint distribution, mainly the heterogeneity variance. Solutions to the integrals in Equation A4 can be obtained by Monte Carlo simulation. The simulated integrals through *N* Monte Carlo draws of v are given by

$$\text{(A5)} \qquad S_{jt}\left(\delta_{jt}|\theta_2\right) \approx \frac{1}{N}\sum_{i=1}^{N} Pr_{ijt}$$

Each consumer has a different price (tax) elasticity for each individual brand. Equations A6 and A7 show the own- and cross-price (tax) elasticities, respectively:

$$\text{(A6)} \qquad \eta_{ijt} = \frac{\partial s_{jt}}{\partial P_{kt}} \cdot \frac{P_{kt}}{s_{jt}} = \begin{cases} \frac{P_{jt}}{s_{jt}} \int \alpha_i\, S_{ijt}(1-S_{ijt}) dP_v(v) & for\ j=k, \\ \frac{-P_{kt}}{s_{jt}} \int \alpha_i\, S_{ijt} S_{ikt} dP_v(v) & otherwise \end{cases}$$

$$\text{(A7)} \qquad \varsigma_{ijt} = \frac{\partial s_{jt}}{\partial SalesTax_{kt}} \cdot \frac{SalesTax_{kt}}{s_{jt}} = \begin{cases} \frac{SalesTax_{jt}}{s_{jt}} \int \gamma_i\, S_{ijt}(1-S_{ijt}) dP_v(v) & for\ j=k, \\ \frac{-SalesTax_{kt}}{s_{jt}} \int \gamma_i\, S_{ijt} S_{ikt} dP_v(v) & otherwise \end{cases}$$

Identification

Prices are potentially correlated with unobserved product characteristics and/or demand shocks. Following the literature, we include several sets of instruments into the estimation to control for price endogeneity and to generate moment conditions to identify coefficients. Instruments that we adopt include cost shifters of producing CSDs, price of sugar, electricity, crude oil, aluminum, and manufacturing wage rates (BLP; Nevo 2001). We collect prices of crude oil and electricity from the U.S. Energy Information Administration. Sources of sugar prices and aluminum prices are from the Economic Research Service, United States Department of Agriculture. Manufacturing wage rates are from the Bureau of Labor Statistics. Furthermore, we interact cost shifters with firm-specific dummies following Villas-Boas (2007). The rational is that different firms may use inputs differently. We also include Hausman-type instruments (i.e., products' own prices in other markets) (Hausman and Taylor 1981). The intuition is that the prices of the same brand in different markets are correlated (due to the common production cost) but are uncorrelated with market specific demand shocks.

We estimate the demand model specified in Equation 1 using a nonlinear generalized methods of moments (GMM) estimator. Following BLP, we use the nested fixed point maximum likelihood algorithm approach to estimate the model parameters. Let IV be the full set of instrumental variables satisfying

(A8) $$g(\delta) = E[IV'\xi] = 0$$

Then, let Φ be the GMM weighting matrix (a consistent estimate of $E[\mathrm{IV}'\xi\xi'\mathrm{IV}]^{-1}$); the estimated parameters can be solved through the following minimization problem:

(A9) $$\hat{\theta} = \underset{\theta}{argmin}(\xi' IV \Phi IV' \xi)$$

References

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (5), 841-90.

Hausman, Jerry A. and William E. Taylor (1981), "Panel Data and Unobservable Individual Effects," *Econometrica: Journal of the Econometric Society*, 1, 1377-98.

Nevo, Aviv (2000), "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategies*, 9 (4), 513-48.

----- (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69 (2), 307-42.

Villas-Boas, Sofia B (2007), "Vertical Relationships Between Manufacturers and Retailers: Inference with Limited Data," *Review of Economic Studies*, 74 (2), 625-52.