Notes on BLP demand estimation using the random-coefficients logit model.

Daria Pus

11/28/2017

Motivation

The BLP Method is a way to estimate demand curves in markets with differentiated products using aggregate data. This framework accommodates consumer heterogeneity, by allowing taste parameters to vary with individual characteristics through random coefficients. The model produces cross price elasticities that are more realistic than the one in regular logit model (see McFaddens famous red-bus/blue-bus example, McFadden (1974)).

Main Features of the BLP Framework:

- Products are bundles of characteristics.
- Preferences are defined on those characteristics.
- Each consumer chooses a bundle that maximizes its utility.
- Consumers have different preferences for different characteristics, and hence make different choices.
- Simulations are used to obtain aggregated demand.

Data needed for estimation:

Aggregate ”Market” Data:

- Longitudinal: one market/store across time
- Cross-sections: multiple markets/stores in a single period of time
- Panel: multiple markets/stores across time

Typical variables used in estimation:

- Aggregate quantity/market shares
- Product prices and characteristics
- Instrumental variables for prices
- Definition of market size
- Distribution of demographics (sometimes)

The Model

Following Nevo(2000b), assume we observe $t = 1, \ldots, T$ markets, each with $i = 1, \ldots, I_t$ consumers. For each such market we observe aggregate quantities, average prices, and product characteristics for J products. Consumer i's utility of consuming product j in market t can be expressed as

$$u_{ijt} = x_{jt}\beta_i - \alpha_i p_{jt} + \xi_j + \Delta\xi_{jt} + \varepsilon_{ijt}$$

where x_{jt} is a K-dimensional vector of observable characteristics of product j in market t, p_{jt} is the price, ξ_j is the mean of the unobserved product characteristics, $\Delta\xi_{jt}$ is the market specific deviation from this mean, and ε_{ijt} is a mean zero stochastic term. α_i and β_i are individual specific coefficients, such that

$$\begin{pmatrix} \beta_i \\ \alpha_i \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix} + \Pi D_i + \Sigma\nu_i$$

where D_i is a $d \times 1$ vector of demographic variables, Π is a $(K+1) \times d$ matrix of coefficients that measure how the taste characteristics vary with demographics, $\nu_i \sim N(0, I_{K+1})$ and Σ is a scaling matrix.

Finally, utility of outside good is

$$u_{i0t} = \xi_0 + \pi_0 D_i + \sigma_0 \nu_{i0} + \varepsilon_{i0t}$$

The mean utility of outside good is not identified , thus it is common to normalize ξ_0 to zero. The coefficients π_0 and σ_0 are not identified separately from coefficients on a constant that are allowed to vary by household. The coefficients on this constant are interpreted as utility parameters of outside good.

Let $\theta = (\theta_1, \theta_2)$ be a vector containing a;ll the parameters of the model. The vector $\theta_1 = (\alpha, \beta)$ contains linear parameters and the vector $\theta_2 = (\Pi, \Sigma, \pi_0, \sigma_0)$ non-linear parameters. Rewrite

$$u_{ijt} = \delta_{jt}(x_{jt}, p_{jt}, \xi_j, \Delta\xi_{jt}; \theta_1) + \mu_{ijt}(x_{jt}, p_{jt}, \nu_i, D_i; \theta_2) + \varepsilon_{ijt}$$

$$\delta_{it} = x_{jt}\beta - \alpha p_{jt} + \xi_j + \Delta\xi_{jt}$$

$$\mu_{ijt} = [p_{jt}, x_{jt}] * (\Pi D_i + \Sigma \nu_i)$$

The utility is now expressed as mean utility δ_{jt} and a mean zero heteroskedastic deviation from that mean $\mu_{ijt} + \varepsilon_{ijt}$ that captures the effects of random coefficients. The estimation exploits this separation to reduce the number of parameters that enter in a non-linear fashion and generate linear moment conditions.

The market share of the jth product is just an integral over the mass of consumers in the region market (or their D_i and ν_i):

$$s_{jt}(\delta_t, \theta_2) = \int_{D_i, \nu_i} \frac{\exp\{\delta_{jt} + \mu_{ijt}(D_i, \nu_i; \theta_2)\}}{\sum_{k=0}^{J} \exp\{\delta_{kt} + \mu_{ikt}(D_i, \nu_i; \theta_2)\}} dF(D_i, \nu_i; \theta_2) \tag{1}$$

Own elasticity

$$\eta_j = \frac{\partial s_j}{\partial p_j} \cdot \frac{p_j}{s_j} = \frac{p_j}{s_j} \int \alpha_i s_{ij}(1 - s_{ij}) dF(D, \nu)$$

Cross elasticity

$$\eta_{jk} = \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = \frac{p_k}{s_j} \int \alpha_i s_{ij} s_{ik} dF(D, \nu)$$

Good Properties:

- Higher priced products more likely purchased by low α_i customers.
- Cross elasticities vary across products

Estimation Algorithm

Step 0:

- Draw ν_i (and D_i if needed) for a set of consumers (say NS=50).
- Initialize mean utility δ_{jt} based on homogenous logit.
- Choose initial values for θ_2.

Step 1:

- For given θ_2 (and initial draws of ν_i and D_i) compute household deviations from mean utility $\mu_{ijt}(x_{jt}, p_{jt}, \nu_i, D_i; \theta_2)$.
- For given mean utility δ_{jt} and θ_2, compute predicted shares. The integral in (1) can be approximated by simulation with

$$s_{jt}(\delta_t, \theta_2) = \frac{1}{NS} \sum_{i=0}^{NS} \frac{\exp\{\delta_{jt} + \mu_{ijt}\}}{\sum_{k=0}^{J} \exp\{\delta_{kt} + \mu_{ikt}\}}$$

Step 2:

- Contraction Mapping : given θ_2, search for δ_t such that market shares computed in *Step 1* are equal to the observed market shares, $s_{jt} = s_{jt}(\delta_t, \theta_2)$. This is a non-linear system of equations that is solved numerically using contraction mapping proposed by BLP (1995)

$$\delta_t^{h+1} = \delta_t^h + \ln s_{jt} - \ln s_{jt}(\delta_t^h, \theta_2)$$

  Iteration continues until $||\delta_t^{h+1} - \delta_t^h||$ is below a specified tolerance level.

Step 3:

- From δ_t, estimate the linear parameters θ_1 using the fact that $\delta_{jt}(s_{jt}, \theta_2) - (x_{jt}\beta - \alpha p_{jt}) = \xi_{jt}$. The IV moment conditions are $E[Z'\xi] = 0$.

$$\hat{\theta}_1 = (X_1' Z W^{-1} Z' X_1)^{-1} X_1' Z W^{-1} Z' \delta(\theta_2)$$

  here X_1 - product characteristics that enter linear part of the estimation, Z - instruments for endogenous variables, W - consistent estimate of $E[Z'\xi\xi' Z]$.

- Compute GMM objective function

$$Q(\theta_2) = \hat{\xi}(\theta_2)' Z W Z' \hat{\xi}(\theta_2)$$

  where $\hat{\xi}(\theta_2)$ are GMM residuals.

Step 4:

- Minimize $Q(\theta_2)$ over θ_2 with *Steps 1-3* nested for every θ_2 trial.

Estimating Standard Errors

The asymptotic covariance matrix for the GMM estimator is given by

$$\hat{V}(\hat{\beta}_{GMM}) = N(\Lambda' Z W Z' \Lambda)^{-1} (\Lambda' Z W \hat{S} Z' \Lambda)(\Lambda' Z W Z' \Lambda)^{-1}$$

Where

- Λ is the Jacobian of the moment conditions with respect to θ, $\Lambda = [X_1, \nabla_{\theta_2}(\hat{\delta}(\theta_2))]$. For the linear parameters θ_1 , this reduces to the corresponding variables X_1, while we need to take the derivatives with respect to θ_2.

- $\hat{S} = \frac{1}{NS}\sum_{i=1}^{NS} Z_i' \hat{\xi}_i \hat{\xi}_i' Z_i$

- W is the weighting matrix used in the GMM estimation

$\nabla_{\theta_2}(\hat{\delta}(\theta_2))$ needs to be computed as an integral over the consumer heterogeneity, and uses simulated individual shares. Since these gradients are not needed in the estimation, it should be computed outside the estimation loops (after θ_2 was estimated).

$$\nabla_{\theta_2}(\hat{\delta}(\theta_2)) = d\delta_t = \begin{bmatrix} \frac{\partial \delta_{1t}}{\partial \theta_{21}} & \cdots & \frac{\partial \delta_{1t}}{\partial \theta_{2L}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \delta_{jt}}{\partial \theta_{21}} & \cdots & \frac{\partial \delta_{Jt}}{\partial \theta_{2L}} \end{bmatrix} = \begin{bmatrix} \frac{\partial s_{1t}}{\partial \delta_{1t}} & \cdots & \frac{\partial s_{1t}}{\partial \delta_{Jt}} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_{Jt}}{\partial \delta_{1t}} & \cdots & \frac{\partial s_{Jt}}{\partial \delta_{Jt}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial s_{1t}}{\partial \theta_{21}} & \cdots & \frac{\partial s_{1t}}{\partial \theta_{2L}} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_{jt}}{\partial \theta_{21}} & \cdots & \frac{\partial s_{Jt}}{\partial \theta_{2L}} \end{bmatrix}$$

$$\frac{\partial s_{jt}}{\partial \delta_{jt}} = \frac{1}{NS}\sum_{i=1}^{NS} s_{ijt}(1 - s_{ijt})$$
$$\frac{\partial s_{jt}}{\partial \delta_{kt}} = \frac{1}{NS}\sum_{i=1}^{NS} s_{ijt}s_{ikt}$$
$$\frac{\partial s_{jt}}{\partial \theta_{2l}} = \frac{1}{NS}\sum_{i=1}^{NS} \nu_i^l s_{ijt}\left(x_{jt}^l - \sum_{k=1}^{J} x_{kt}^l s_{ikt}\right)$$

Standard errors are obtained from the square root of the diagonal of $\hat{V}(\hat{\beta}_{GMM})$.

References

McFadden, Daniel. 1974. “Conditional Logit Analysis of Qualitative Choice Behavior ” Pp. 105-42 in *Frontiers of Econometrics*, edited by P. Zarembka. New York: Academic Press.

Nevo, A. (2000). A practitioner’s guide to estimation of randomcoefficients logit models of demand. *Journal of economics & management strategy*, 9(4), 513-548.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841-890.