Demand in Differentiated-Product Markets – a review

Spring 2009

Finishing up differentiated-product analysis

1. Identification
2. Review of BLP – outline of estimation
    - nested fixed-point routine
    - variation in the data
    - instruments
3. Examples of BLP

Identification

- In these discrete choice models, identification means for every vector of parameters there is a unique vector of choice probabilities.
- If two different vectors of parameters lead to the same vector of choice probabilities, then can't distinguish the true underlying model
- With logit, have a clear mapping:
- Same logic with BLP, just more complicated
- identification relies on parametric assumptions (e.g. T1EV for iid error term). The idea is that the parametric forms are flexible, and so can approximate true underlying indirect utility function.
- Not that much work done on pushing towards nonparametric identification (mention a few papers at end)

Example: Logit case

One simple case of inversion step is the MNL case:

$$\hat{s}_j(\delta_1, \ldots, \delta_J) = \frac{\exp(\delta_j)}{1 + \sum_{k=1}^{J} \exp(\delta_k)} \tag{1}$$

The system of equations for matching predicted and actual is (after taking logs)

$$\log(s_1) = \delta_1 - \log(1 + \sum_{k=1}^{J} \exp(\delta_k)) \tag{2}$$

$$\vdots \tag{3}$$

$$\log(s_J) = \delta_J - \log(1 + \sum_{k=1}^{J} \exp(\delta_k)) \tag{4}$$

$$\log(s_0) = \delta_0 - \log(1 + \sum_{k=1}^{J} \exp(\delta_k)) \tag{5}$$

which gives us

$$\delta_j = \log(s_j) - \log(s_0) \tag{6}$$

So in the second step, run an IV regression

$$\log(s_j) - \log(s_0) = X_j\beta - \alpha p_j + \xi_j \tag{7}$$

To calculate s_0 need to make an assumption about the total size of the market! Not always innocuous.

BLP routine

indirect utility function, person i, product j

$$u_{ij} = X_j\beta + \xi_j - \alpha p_j + \sum_k \sigma_k \nu_{ik} x_{jk} + \epsilon_{ij} \tag{8}$$

where x_{jk} are elements of the vector X_j.

Can separate the parameters to be estimated in two groups, linear parameters $\theta_1 = \{\beta, \alpha\}$ and non-linear parameters $\theta_2 = \sigma$.

Estimation routine

1. Before estimation – pick # of consumer (e.g. $N = 50$) and draw their individual tastes, ν_{ij}.

2. Outer loop – searching over θ_2 to minimize the criterion function (i.e. the moments). So pick an initial guess of non-linear parameters, $\hat{\theta}_2$, and compute $\mu(\theta_2) = \sum_k \sigma_k \nu_{ik} x_{jk}$.

3. Inner loop – find δ_j such that model prediction of shares exactly match data solve for θ_1

   (a) pick an initial guess of δ_j

   (b) compute (via simulation)

$$\hat{s}_j = \int \frac{\exp(\delta_j + \mu_{ij})}{1 + \sum_h \exp(\delta_h + \mu_{ih})} dF \tag{9}$$

$$\hat{s}_j = \frac{1}{N} \sum_{i=1}^{N} \frac{\exp(\delta_j + \mu_{ij})}{1 + \sum_h \exp(\delta_h + \mu_{ih})} \tag{10}$$

   (c) update δ via:

$$\delta' = \delta + \log(s) - \log(\hat{s}(\delta, \theta_2)) \tag{11}$$

   (d) recompute predicted market shares, repeat until converge to a *unique* δ (proof relies on Brouwer's fixed-point theorem) that equates predict with actual market share. *This is an important part of identification*

4. With $\hat{\delta}$ run 2SLS to compute θ_1 and infer ξ. Calculate the criterion

$$G(\theta_1, \theta_2)' W G(\theta_1, \theta_2) \tag{12}$$

where W is a weighting matrix. (More work if also computing the supply side — but fits in here)

5. Go back to the outer loop, choose another θ_2, etc.

weighting matrix

- Normally use a 2 step procedure
- In the first stage, set W equal to the identity matrix. So every moment is equally weighted, and just squaring and summing across moments. Label resulting parameters which minimize moments θ_I and the unobserved characteristics as ξ_I.
- In second stage, use the first stage results to compute the optimal weighting matrix.

$$A = E[Z' \xi_I \xi_I' Z] \tag{13}$$
$$W = A^{-1} \tag{14}$$

Note – sometimes can't invert the full matrix – then just use variance terms.

There is a large literature on optimal weighting matrices — above the "standard" method, but there are other techniques.

Variation in the data

- What kind of variation in the data will allow for precise estimates?
- Need to see repeated cross-sections with variation in market shares
- For example, BLP (1995) has around 10 years, where goods entered and exited. Entry and exit really generates the necessary variation. Only having price change sometimes is not enough, b/c relative prices may not change much. Recall this is an age-old problem. E.g. work on consumer choice of banks has run into problems b/c interest rates on deposits hardly vary over time or across regions.

- Having both a demand and supply side generates smaller standard errors – the extra structure on the problem helps estimation (as opposed to only estimating the demand side).

- As with all empirical work, start simple (regular logit) and slowly build in complexity. In standard errors are too large with regular logit, hard to see why they would shrink with complexity.

- Linear parameters are easy to estimate, but adding non-linear parameters can dramatically increase the search time in the outer loop.

Instruments

- Problem: firms & consumers observe ξ and so expect ξ and price to be correlated. But the moment condition is that $E[\xi|X] = 0$.

- See Nevo (2001) for a good discussion

- Usual instruments are other products characteristics

- $\text{char}_j \to p_j \to p_k$ so char_j should be correlated with p_k.

- Assumed with characteristics are exogenous, so no correlation with char_j and ξ_k.

- People usually fine with these instruments, but do have to worry about this assumption is looking at longer periods of time or products which evolve rapidly.

- In practice, use aggregates of other products' characteristics (e.g. mean of all other firms' products characteristics).

On the frontier . . .

1. "A Simple Nonparametric Estimator for the Distribution of Random Coefficients in Discrete Choice Models" by Patrick Bajari, Jeremy Fox, Kyoo il Kim & Stephen Ryan

   Getting away from the parametric identification of these BLP-style models (and other random coeff. models). Addressing the usual practice of assuming σ is normally distributed.

   Also some working papers by Steve Berry and Elie Tamer.

2. "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation" Jean-Pierre Dube, Jeremy Fox & Che-Lin Su

   The double loop for BLP introduces complexity to the code, which increases the probability of errors and so the probability of incorrect results. They introduce a new estimating technique – minimization with constraints.

   Nested fixed-point

   $$\min_{\theta} g(s^{-1}(S;\theta))' W g(s^{-1}(S;\theta)) \tag{15}$$

   New approach

   $$\min_{\theta,\xi} g(\xi)' W g(\xi) \tag{16}$$

   $$\text{subject to: } s(\xi,\theta) = S \tag{17}$$

3. Dynamic consumer demand models – we'll see later in the semester.