Extensions and Variants

Chris Conlon

September 22, 2020

Grad IO

BLP Extensions: Demographics

- It is helpful to allow for interactions with consumer demographics (such as income).
- A few ways to do this:
  - You could just use cross sectional variation in s_{jt} and $\overline{y}_t$ (mean or median income).
  - Better: Divide up your data into additional "markets" by demographics: do you observe $\mathfrak{s}_{jt}$ at this level? [May not be possible!]
  - Better: Draw y_{it} from a geographic specific income distribution. Draw ν_i from a general distribution of unobserved heterogeneity.
- Ex: Nevo (2000) Cereal demand sampled individual level D_i from geographic specific CPS data
- Joint distribution of income, income-squared, age, child at home.

$$\beta_i = \overline{\beta} + \Pi D_i + \sigma \nu_i$$

BLP Extensions: Panel Data

- with enough observations on the same product it is possible to include fixed effects

$$\delta_{jt}(\widetilde{\theta}_2) = x_{jt}\beta - \alpha p_{jt} + \underbrace{\xi_{jt}}_{\xi_j + \xi_t + \Delta\xi_{jt}}$$

- What does ξ_j mean in this context?
- What would ξ_t mean in this context?
- $\Delta\xi_{jt}$ is now the structural error term, this changes our identification strategy a little.
- We need instruments that change within product and across market.
  - ie: $z_{jt} - \overline{z}_{\cdot t} - \overline{z}_{j\cdot} = \Delta z_{jt}$ has to have some variation left!

Extensions: Micro Data (Petrin 2002), (microBLP 2004)

Suppose we had additional data on behavior of individuals (in addition to aggregate market shares).

- Examples:
    - For some customers have answer to "Which car would you have purchased if the car you bought was not available?"
    - Demographic data on purchasers of a single brand.
    - Full individual demographic and choice data.

Extensions: Micro Data: Nielsen Panelists

Nielsen data surveys panelists on everything they buy with a UPC code including what store they purchased from.

- Also tracks household characteristics (Race, Income, Education, HH Size, etc.)
- Can calculate covariance of characteristics (such as price) with demographics (income, education, etc.) conditional on purchase
- Can calculate purchase probability conditional on demographics: Did you buy any yogurt this trip, week, month, year?

Should we use these as individual data? Or Aggregate data from scanner data with additional moments?

Extensions: Micro Data (Petrin 2002), (microBLP 2004)

- Previously we had moment conditions from orthogonality of structural error (ξ) and (X, Z) in order to form our GMM objective.

$$E[\xi_{jt}|z_{jt}] = 0 \rightarrow E[\xi'_{jt} Z_{jt}] = 0$$

- We can incorporate additional information using "micro-moments" or additional moment conditions to match the micro data.
  - $Pr($ i buys j $|y_i \in [0, \$20K]) = c_1$ or $Cov(d_i, s_{ijt}) = c_2$
  - Construct an additional error term ζ_1, ζ_2 and interact that with instruments to form additional moment conditions.
  - Econometrics get tricky when we have a different number of observations for $E[\zeta' Z_m] = 0$ and $E[\xi' Z_d] = 0$.
    - May not be able to get covariance of moments taken over different sets of observations!
    - People often assume optimal weight matrices are block diagonal.

Alternative: Vertical Model (Bresnahan 1987)

- Imagine everyone agreed on the quality of the products offered for sale.
- The only thing people disagree on is willingness to pay for quality

$$U_{ij} = \overline{u} + \delta_j - \alpha_i p_j$$

- How do we estimate?
  - Sort goods from $p_1 < p_2 < p_3 \ldots < p_J$.
    It must be that $\delta_1 < \delta_2 < \ldots < \delta_J$. Why?
  - Normalize o.g. to 0 so that $0 > \delta_1 - \alpha_i p_1$ or $\alpha_i > \delta_1 / p_1$.
  - $s_0 = F(\infty) - F(\frac{\delta_1}{p_1}) = 1 - F(\frac{\delta_1}{p_1})$ where $F(\cdot)$ is CDF of α_i.
  - In general choose j IFF:

$$\frac{\delta_{j+1} - \delta_j}{p_{j+1} - p_j} < \alpha_i < \frac{\delta_j - \delta_{j-1}}{p_j - p_{j-1}}$$
$$s_j = F\left(\frac{\delta_{j+1} - \delta_j}{p_{j+1} - p_j}\right) - F\left(\frac{\delta_j - \delta_{j-1}}{p_j - p_{j-1}}\right)$$

Alternative: Vertical Model (Bresnahan 1987)

Estimation

- Choose parameters θ of $F(\cdot)$ in order to best match s_j.
  - Can do MLE $\arg\max_\theta \sum_j -\mathfrak{s}_j \log s_j(\theta)$.
  - Can do least squares $\sum_j (\mathfrak{s}_j - s_j(\theta))^2$.
  - Can do IV/GMM if I have an instrument for price. $\delta_j = x_j \beta + \xi_j$.
  - Extremely easy when $F \sim \exp(\lambda)$.
- What about elasticities?
  - When I change the price of j it can only affect (s_{j-1}, s_j, s_{j+1}).
  - We have set all of the other cross-price elasticities to be zero.
  - If a luxury car and a truck have similar prices, this can create strange substitution patterns.

Pure Characteristics Model: Berry Pakes (2001/2007)

$$u_{ij} = \delta_j + \beta_i x_{jt} + \xi_{jt} + \underbrace{\sigma_e}_{\to 0} \cdot \varepsilon_{ijt}$$

- Can think of this like random coefficients model where we take the variance of ϵ to zero.
- Can think of this a vertical model, with vertical tastes over several characteristics.
    - PCs: everyone prefers more Mhz, more RAM, and more storage but differ in WTP.
    - Possible that there is no PC specific ε.
- Advantages
    - Logit error means there is always some substitution to all other goods.
    - Reality may be you only compete with a small number of competitors.
    - Allows for crowding in the product space.
- Disadvantage: no closed form for s_j, so estimation is extremely difficult.
- Minjae Song (Homotopy) and Che-Lin Su (MPCC) have made progress using two different approaches.

Even More Flexibility (Fox, Kim, Ryan, Bajari)

Suppose we wanted to nonparametrically estimate $f(\beta_i|\theta)$ instead of assuming that it is normal or log-normal.

$$s_{ij} \quad = \quad \int \frac{\exp[x_j \beta_i]}{1 + \sum_k \exp[x_k \beta_i]} f(\beta_i|\theta)$$

- Choose a distribution $g(\beta_i)$ that is more spread out that $f(\beta_i|\theta)$
- Draw several β_s from that distribution (maybe 500-1000).
- Compute $\hat{s}_{ij}(\beta_s)$ for each draw of β_s and each j.
- Holding $\hat{s}_{ij}(\beta_s)$ fixed, look for w_s that solve

$$\min_w \left(s_j - \sum_{s=1}^{ns} w_s \hat{s}_{ij}(\beta_s) \right)^2 \quad \text{s.t. } \sum_{s=1}^{ns} w_s = 1, \quad w_s \geq 0 \quad \forall s$$

Even More Flexibility (Fox, Kim, Ryan, Bajari)

- Like other semi-/non- parametric estimators, when it works it is both flexible and very easy.
- We are solving a least squares problem with constraints: positive coefficients, coefficients sum to 1.
- It tends to produce sparse models with only a small number of β_s getting positive weights.
    - Why? There is an L_1 penalty term (We are doing non negative LASSO!)
- This is way easier than solving a random coefficients logit model with all but the simplest distributions.
- There is a bias-variance tradeoff in choosing $g(\beta_i)$.
- Incorporating parameters that are not random coefficients loses some of the simplicity.
- I have no idea how to do this with large numbers of fixed effects.

Fully Nonparametric Demand (Compiani 2019)

Takes identification arguments in Berry Haile (2014) to the data. Looks at a sieve approximation to

$$D_{jt}^{-1}(\mathcal{S}_t, \widetilde{\theta}_2)$$

Using the Bernstein Polynomials

- Bernstein polynomials make it possible to enforce shape restrictions and monotonicity which is important
- Estimates demand for strawberries (organic vs. non-organic)
- Suggests that both for markups and merger effects we don't have sufficiently flexible demand models.