

Advanced Applied Econometrics

Static discrete choice with market-level data

Hannes Ullrich

DIW Berlin and University of Copenhagen

DIW Office 5.2.020, email: hullrich@diw.de

20 June 2025, DIW

Organization

- ▶ Next class on June 25, 15:30-18:30
 - ▶ Single agent dynamic discrete choice
 - ▶ Read Rust (1987): "Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher"
 - ▶ Graded problem set will be distributed this week. Due by Thursday, July 6, 23:59.

Plan for today

- ▶ Recap BLP: The Random Coefficient Logit Model of Demand
- ▶ Go through BLP code
- ▶ Discuss addition of supply-side moments

Recap BLP

- ▶ Individual choice model using market-level data with
 - ▶ horizontal (ϵ, μ) and vertical (δ, ξ) product differentiation
 - ▶ (price) endogeneity
 - ▶ Isolate mean utility to estimate β coefficients by linear IV estimation
 - ▶ unobserved consumer heterogeneity
 - ▶ flexible substitution patterns
 - ▶ In homogenous logit, only market shares matter
 - ▶ In heterogeneous logit, closeness in characteristics space \rightarrow product differentiation matters
 - ▶ static pricing game on supply side can improve identification
 - ▶ quasilinear preferences allow computing changes in consumer welfare

Recap BLP: Identification

- ▶ exogenous variation in observed characteristics
- ▶ choice set variation (if data on multiple markets available)
- ▶ variation in preferences \rightarrow variation in choice probabilities across consumer groups
- ▶ formal positive nonparametric identification results by Berry and Haile (Ecta 2014, ARE 2016), Fox and Gandhi (Rand 2016), Fox, Kim, Ryan, and Bajari (JoE 2012)
 \rightarrow functional forms and distributional assumption not necessary for identification, standard IV conditions are sufficient
- ▶ Intuition for BLP instruments: assuming $E[\xi_j Z_j] = 0$ shuts off opportunity for model to explain data by systematic correlation between ξ_j and observable local market structure.
- ▶ Moment conditions with such IVs, in principle, allow identification of $\{\sigma, \delta, \alpha, \beta\}$

BLP widely used but problems

► Econometric

- Identifying unobserved heterogeneity with aggregate data can be hard in practice, Petrin (JPE 2002), BLP (JPE 2004) add microdata
- Weak instrumental variables due to lack of cost shifters - Armstrong (Ecta 2016), Reynaert and Verboven (JoE 2014), Berry and Haile (2014), Gandhi and Houde (2019), Gandhi and Houde (2020), handbook chapter Gandhi and Nevo (2021)
- "Logit" assumption / Welfare analysis with many products, entry/exit, Akerberg and Rysman (Rand 2005), Berry and Pakes (IER 2007)
- Measurement error in market shares (small T , large J), moment inequalities, Gandhi, Lu, and Shi (QE 2023)

BLP widely used but problems

- ▶ Numerical
 - ▶ Error tolerance in inner loop δ , premature 'convergence', speed of convergence
Dubé, Fox, and Su (Ecta 2012), Knittel and Metaxoglu (ReStat 2012), Reynaerts, Varadhan, and Nash (2012)
 - ▶ Quality of solvers in estimation, importance of starting values,
Dubé, Fox, and Su (Ecta 2012), Knittel and Metaxoglu (ReStats 2014)
 - ▶ Numerical integration techniques and simulation error,
Judd and Skrainka (2011), Chiou and Walker (JoE 2007)
- ▶ Implications for elasticity estimates and, in consequence, for measure of market power, merger evaluation, welfare gains from new products/technologies, ...?
- ▶ Solutions to most of these problems implemented in Conlon and Gortmaker (2020):
<https://github.com/jeffgortmaker/pyblp>

Nested fixed point algorithm (BLP 1995)

Functions to be called (directly and indirectly) from main script

- ▶ Contraction mapping: $\delta^{h+1} = \delta^h + \ln(S) - \ln(s(\delta^h, \sigma))$
- ▶ Market shares: $s_{jt}(\delta_t, \sigma) \approx \sum_{i=1}^n \phi_i \frac{\exp(\delta_{jt} + \mu_{jt}(v_i))}{1 + \sum_{l=1}^J \exp(\delta_{lt} + \mu_{lt}(v_l))}$
- ▶ Market share derivatives Δ_t :

$$\sum_{i=1}^n \phi_i [-\alpha s_{ijt}(1 - s_{ijt})] \quad \forall j = k, \quad \sum_{i=1}^n \phi_i [\alpha s_{ijt}s_{ikt}] \quad \forall j \neq k$$

Nested fixed point algorithm (BLP 1995)

Functions to be called (directly and indirectly) from main script

► Contraction mapping: $\delta^{h+1} = \delta^h + \ln(S) - \ln(s(\delta^h, \sigma))$

► Market shares: $s_{jt}(\delta_t, \sigma) \approx \sum_{i=1}^n \phi_i \frac{\exp(\delta_{jt} + \mu_{jt}(v_i))}{1 + \sum_{l=1}^J \exp(\delta_{lt} + \mu_{lt}(v_i))}$

► Market share derivatives Δ_t :

$$\sum_{i=1}^n \phi_i [-\alpha s_{ijt}(1 - s_{ijt})] \quad \forall j = k, \quad \sum_{i=1}^n \phi_i [\alpha s_{ijt}s_{ikt}] \quad \forall j \neq k$$

► For data generation, need also equilibrium prices, using FOC: $0 = c_t - p_t - \Delta_t^{-1} s_t$

Nested fixed point algorithm (BLP 1995)

- ▶ GMM objective function, minimization problem based on moment conditions $E[g_{jt}(z_{jt})\zeta_{jt}] = 0$:

$$\min_{\theta} \zeta(\theta)' g(z)' A g(z) \zeta(\theta)$$

- ▶ For simplicity, no analytical gradients. In practice, use if possible!

BLP Code

Notes on numerical integration

Stochastic / Monte Carlo

- ▶ “Monte Carlo is the art of approximating an expectation by the sample mean of a function of simulated random variables” - Statistical Genetics lecture notes, UC Berkeley.
- ▶ Pseudo-random - standard random number generator
- ▶ Quasi-random - more uniform coverage, e.g. Halton, modified Latin hypercube sampling

Notes on numerical integration

Non-stochastic / Quadrature

- ▶ Gaussian Hermite product rule
 - ▶ Difficult for high-dimensional distributions and complicated integrands
- ▶ Sparse grid integration (Heiss and Winschel, JoE 2008)
 - ▶ Subset of nodes from product rule
- ▶ Simulation bias in MSL and MSS (ln in simulated $\ln P_n(\theta)$ is a nonlinear transformation), not in MSM

Supply-side equilibrium Model

- ▶ How are prices determined in oligopolistic markets with differentiated products?
- ▶ → counterfactual policy analysis
- ▶ Increased efficiency of demand estimates by adding structure on the supply side
- ▶ Assume prices are result of firms' optimization problem
- ▶ Use information from first-order conditions for estimation
- ▶ Single-product vs. multi-product firms
- ▶ Specify cost function and equilibrium notion

Supply-side equilibrium Model

Berry (1994) single-product firms

- ▶ Profits of firm j are given by $\Pi_j(p) = p_j q_j(p) - C_j(q_j(p))$
- ▶ Conduct assumption: compete à la Nash in prices
- ▶ The first-order condition is:

$$q_j + (p_j - mc_j) \frac{\partial q_j}{\partial p_j} = 0$$

- ▶ which can be rewritten as

$$p_j = mc_j + b_j(p)$$

where $b_j(p) = \frac{-q_j}{\partial q_j / \partial p_j}$ is the *markup*

Multiproduct Firms

BLP (1995) look at the more general case of a multiproduct firm f producing a subset \mathfrak{S}_f of the J products in the market

- ▶ Profits of firm f are given by:

$$\Pi_f = \sum_{j \in \mathfrak{S}_f} (p_j - mc_j) Ms_j(p, x, \theta_1, \theta_2)$$

- ▶ The usual assumption is that firms compete à la Nash in prices

Multiproduct Firms

- ▶ The first-order condition is:

$$s_j(p, x, \theta_1, \theta_2) + \sum_{r \in \mathfrak{S}_f} (p_r - mc_r) \frac{\partial s_r(p, x, \theta_1, \theta_2)}{\partial p_j} = 0$$

- ▶ which can be rewritten in vector notation as:

$$p_j = mc_j + \Delta(p, x, \theta_1, \theta_2)^{-1} s(p, x, \theta_1, \theta_2)$$

where the (j, r) element of the J by J matrix $\Delta(p, x, \theta_1, \theta_2)$ is

$$\Delta_{jr} \begin{cases} 1 \times \frac{\partial s_r}{\partial p_j} = \frac{\partial s_r}{\partial p_j} & \text{if } r \text{ and } j \text{ are produced by the same firm} \\ 0 \times \frac{\partial s_r}{\partial p_j} = 0 & \text{otherwise} \end{cases}$$

Multiproduct Firms: note on conduct

- Rewrite profit maximization problem:

$$\Pi_f = \sum_{j \in \mathfrak{S}_f} (p_j - mc_j) Ms_j(p, x, \theta_1, \theta_2) + \kappa_{fg} \sum_{j \in \mathfrak{S}_g} (p_j - mc_j) Ms_j(p, x, \theta_1, \theta_2)$$

- The first-order condition is then:

$$s_j(p, x, \theta_1, \theta_2) + \sum_{r \in (\mathfrak{S}_f, \mathfrak{S}_g)} \kappa_{fg} (p_r - mc_r) \frac{\partial s_r(p, x, \theta_1, \theta_2)}{\partial p_j} = 0$$

- Instead of 0's and 1's, now $\kappa_{fg} \in [0, 1]$ represents how much firm f cares about profits of g
- In reality, $\kappa_{fg} \in (0, 1)$, but evidence that $\kappa_{fg} > 0$ not necessarily “anti-competitive” – just deviation from static Bertrand pricing
- Berry and Haile (2014): κ_{fg} can be identified using IV that shifts demand but not supply (“rotate demand”)

Multiproduct Firms

- ▶ For simplicity, assume that marginal costs are constant and take the form:

$$\ln(mc) = w_j\gamma + \omega_j \quad (1)$$

- ▶ The equation to be estimated is therefore

$$\ln(p - b(p, x, \theta_1, \theta_2)) = w_j\gamma + \omega_j$$

where w_j are observed and ω_j unobserved product characteristics

- ▶ The markup $b(p, x, \theta_1, \theta_2) = \Delta(p, x, \theta_1, \theta_2)^{-1}s(p, x, \theta_1, \theta_2)$ depends on demand parameter and, through the equilibrium price, on the unobserved cost component ω

GMM estimation

- ▶ Estimate demand and supply simultaneously. Define additional instruments $z_j = (x_j, w_j)$.
- ▶ Additional moment conditions

$$E[\xi_j|z] = E[\omega_j|z] = 0$$

where $\omega = mc - [x, w]\gamma$