ECON-L1350 - Empirical Industrial Organization PhD I: Static Models

Lecture 4

Tanja Saxell

Part 2: BLP

- Computationally demanding, progress on these challenges
- Many variations of the model in the literature

Part 2: Random Coefficient Models / BLP

Part 2: Random Coefficient Models / BLP

- Computational challenges
- Best practices
- Extensions such as micro data

Part 2: Key Papers

- Conlon, C. and Gortmaker, J. 2020. Best practices for differentiated products demand estimation with PyBLP. The RAND Journal of Economics, 51: 1108-1161.
- Dubé, J.-P., Fox, J.T. and Su, C.-L. 2012. Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation. Econometrica, 80: 2231-2267.
- Berry, S., and P. Haile. 2021. "Chapter 1 - Foundations of demand estimation," The Handbook of Industrial Organization, Editor(s): Kate Ho, Ali Hortaçsu, Alessandro Lizzeri, Elsevier, Volume 4, Issue 1, 2021, 1-62. (Chapter 6)
- Petrin, A. 2002. Quantifying the Benefits of New Products: The Case of the Minivan. Journal of Political Economy, 110(4).

Computational challenges and best practices

BLP Algorithm: Reminder

$$u_{ijt} = \underbrace{x_{jt}\beta - \alpha p_{jt} + \xi_{jt}}_{\delta_{jt}} + x_{jt}\sigma v_{it} + e_{ijt}$$

- Outer loop: search over trial values of the parameter vector $\theta = (\theta_1, \theta_2)$, with $\theta_1 = (\beta, \alpha)$ (linear parameters) $\theta_2 = \sigma$ (nonlinear parameters)
- Inner loop: given θ, find a solution for $\delta_t(\theta_2) \equiv \delta_j(s_t, x_t, \theta_2)$ in each market t such that $s_{jt} = s_j(\delta_t, x_t, \theta)$ as fixed point iteration
- Then calculate $\xi_{jt} \equiv \delta_{jt}(\theta_2) - (x_{jt}\beta - \alpha p_{jt})$

```
begin outer loop
try new θ
    begin inner loop
     solve contraction mapping (fixed point iteration)
    end inner loop
calculate GMM criterion
end outer loop
```

BLP Challenges

- The BLP has a unique fixed point, and the fixed-point iteration is guaranteed to converge from any starting point
- However, the BLP contraction mapping $\delta_t^{(k)} = \delta_t^{(k-1)} + \left[\ln(s_t) - \ln(s(\delta_t^{(k-1)}, x_t, \theta_2))\right]$ converges only linearly and can prove to be a time-consuming procedure, especially in applications with a large number of observations
  - If sequence $\{x^k\}$ converges to x^* and there exists $\lambda > 0$ and $q \geq 1$ s.t. $\lim_{n\to\infty} \frac{|x^{k+1}-x^*|}{|x^k-x^*|^q} = \lambda$ we say that q is the rate of convergence
  - q represents how quickly the sequence approaches its limit
  - $q = 1$: linear, $q = 2$: quadratic etc., ordered from slowest to fastest.

BLP: Challenges (e.g. Conlon and Gortmaker, 2020)

- Inner loop can be slow
- Outer loop optimization is hard
- Involves a nonlinear change of variables from the space of observed market shares to the space of mean utilities for products
- Parameters governing the nonlinear change of variables are unknown
- This results in a non-linear, non-convex optimization problem with a simulated objective function
- Optimization routines such as Nelder-Mead (e.g. Matlab's fminsearch) can fail (convergence to a local minimum)
  - Recommendation: consider a number of different starting values and optimization routines.

BLP: Challenges

The main challenge of the BLP nested fixed point algorithm is solving the system of market shares $s_t = s(\delta_t, x_t, \theta_2)$:

$$\delta_t^{(k)} = \delta_t^{(k-1)} + \left[\ln(s_t) - \ln(s(\delta_t^{(k-1)}, x_t, \theta_2))\right]$$

- Although mathematically there is a unique solution, it is impossible, numerically speaking, to choose a vector δ_t that solves $s_{jt} = s(\delta_t, \theta_2)$ exactly.
- Instead, we must solve the system of equations to some tolerance.

BLP: Challenges

- Tolerances are important, with less stringent stopping criterion for the inner loop can get wrong answers (e.g. Dube, Fox, and Su, 2012)
- The inner loop error propagates into the outer loop GMM objective function and its derivatives, which may cause an optimization routine to fail to converge
- It is also possible to set a tolerance which is too tight and thus can never be satisfied (Conlon and Gortmaker, 2020).
    - They prefer to set the tolerance between 1E-14 and 1E-12 as the machine epsilon or detectable difference between two double precision floating point numbers is around 1E-16.

Accelerating Contraction (Conlon and Gortmaker, 2020)

- A direct approach would be to solve a system of equations $s_t = s(\delta_t, x_t, \theta_2)$ using Newton-type methods.
- Newton's method is a powerful technique - in general the convergence is quadratic.
- Idea of Newton's method for solving a system of equations $f(x) = 0$:
    - Using basic calculus, the linear approximation of f at x_0 is $f(x) \approx f(x_0) + \Psi(f(x_0))(x - x_0)$ where $\Psi(f(x_0))$ is a matrix of partial derivatives (Jacobian matrix) evaluated at x_0.
    - To get new x_1, we can solve this equation by $x_1 = x_0 - \Psi^{-1}(f(x_0))f(x_0)$.
    - Iterate until convergence, $(x_k - x_{k-1}) < \epsilon_{tol}$.
    - Matlab: fsolve

Accelerating Contraction (Conlon and Gortmaker, 2020)

Conlon and Gortmaker propose the following Newton (also known as Newton-Raphson) iteration for solving $s_t = s(\delta_t, x_t, \theta_2)$ for step size λ

$$\delta_t^{(k)} = \delta_t^{(k-1)} - \lambda \Psi^{-1}(\delta_t^{(k-1)}, x_t, \theta_2) s(\delta_t^{k-1}, x_t, \theta_2)$$

- Each Newton-Raphson iteration would require computation of:
- A vector of market shares $s_t(\cdot)$
- The Jacobian matrix $\Psi_t(\cdot)$ as well as its inverse (can be costly, ways to speed up).
    - Variants of the algorithm generally involve modifying the step-size λ or approximating $\Psi^{-1}(\cdot)$ in ways that avoid calculating the inverse at each step.

BLP: Challenges

- BLP algorithm also forces constraints on market shares and linear parameters hold exactly at every guess of θ_2 when we care only about constraints' holding at the final solution.
- Progress also on this challenge.

Constrained Optimization

- E.g. Dube, Fox, and Su (2012): alternative computation approach for the same BLP estimator.
- They propose a routine to solve a constrained optimization problem: minimize GMM objective function over parameters, subject to constraint that the inner loop fixed point equations hold.
- Single-step optimization, no inner/outer loop.

DFS Algorithm: Basic Idea

Constrained optimization algorithm of BLP estimator:

$$\begin{aligned} \min_{\theta,\xi} \quad & \xi' z' W z' \xi \quad \text{s.t.} \\ log(s_{jt}) \quad &= \quad log(s_j(x_t, \xi_t, \theta)) \ \forall j, t \\ \text{where } s_j(x_t, \xi_t, \theta) \quad &= \quad \frac{1}{NS} \sum_{i=1}^{NS} \frac{exp(x_{jt}\beta + \xi_{jt} + x_{jt}\tilde{\beta}_{it})}{1 + \sum_k exp(x_{kt}\beta + \xi_{kt} + x_{kt}\tilde{\beta}_{it})} \end{aligned} \tag{1}$$

Note that here ξ_{jt} are parameters whose values, together with θ, must equate predicted and observed shares.

- Note: a large number of parameters!

DFS vs. BLP

- DFS: Solving a constrained optimization problem w.r.t. θ and ξ

$$\begin{aligned} \min_{\theta,\xi} \quad & \xi' z' W z' \xi \quad \text{s.t.} \\ log(s_{jt}) &= log(s_j(x_t, \xi_t, \theta)) \ \forall j, t \end{aligned}$$

  The algorithm is a mathematical program with equilibrium constraints (MPEC)

- BLP: Nesting a contraction mapping for each trial value of θ

$$\min_{\theta} \quad \xi(\theta)' z' W z' \xi(\theta)$$

  A nested fixed point (NFP) algorithm.

DFS vs. BLP

- BLP and DFS define different algorithms to produce the same statistical estimator.
  - $\hat{\theta}_{NFP} \approx \hat{\theta}_{MPEC}$ but for numerical differences in the optimization routine.
- A choice of algorithm should mostly be about computational convenience, both can work well when carefully executed!
  - By eliminating the nested calls to a contraction mapping, MPEC can be faster than NFP
  - On the other hand, MPEC has a large number of auxiliary parameters to be estimated
- Current best practice of BLP, with several improvements to the original version: pyBLP (Python) implementation of Conlon and Gortmaker (2019)

Extensions: demographics, panel data, micro data

BLP with Demographics (and Aggregate Data)

- Can also allow consumer preferences to vary as a function of the individual characteristics such as income and age (already in the original BLP)
- A few ways to do this:
  - 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 s_{jt} at this level?
  - Better: Draw y_{it} from a geographic specific income distribution. Draw ν_i from a general distribution of unobserved heterogeneity.

BLP with Demographics (and Aggregate Data)

- Example: Nevo (2000): the distribution of consumer taste parameters as a function of demographics d:

$$\beta_{it} = \beta + \Pi d_{it} + \sigma \nu_{it}, \; d_{it} \sim F_d, \nu_{it} \sim F_\nu$$

- Given that no individual data is observed in BLP/Nevo (2000), neither component of the individual characteristics is directly observed in the choice data set.
- We know something about the distribution of the demographics:
  - a nonparametric distribution estimated from other data sources
  - a parametric distribution with the parameters estimated elsewhere, e.g. when the mean and standard deviation (μ_d and σ_d) are known, one alternative is to assume $d_i \sim N(\mu_d, \sigma_d)$

Market Share with Demographics

- The market share of product j is almost as before:

$$\begin{aligned} s_{jt} &= P(a_{it} = j) = \int_{A_{jt}} dF(v, e, d) \\ &= \int_{A_{jt}} dF_v(v) dF_e(e) dF_d(d) \text{ (independence assumption).} \end{aligned}$$

- With Type 1 extreme value distributed error terms (e), this can be approximated by

$$s_j(\delta_t, x_t, \theta_2) \approx \frac{1}{NS} \sum_{i=1}^{NS} \frac{\exp[\delta_{jt}(\theta_2) + x_{jt}(\Pi d_i + \sigma \nu_i)]}{1 + \sum_k \exp[\delta_{kt}(\theta_2) + x_{kt}(\Pi d_i + \sigma \nu_i)]}$$

where d_i and ν_i and draws from the empirical counterparts of F_v and F_d

Different Types of Data

- Aggregate data: market and product characteristics, prices, shares (cross-sectional, panel)
- Micro data: individual characteristics matched to individual consumer choices
- Micro panel data: repeated choices of same consumer over time matched with characteristics (not today)
- Question: Even if we had data on individual consumer choices, should we estimate models using the individual level data? Or aggregate data from the individual level using the BLP type framework (possible with additional moments)?

Micro Data

- Individual consumer choices *not* matched consumer characteristics has no new information relative to market level data
- Consumer characteristics not matched to individual choices is just a version of market level data

Why Micro Data?

- Berry, S., and P. Haile. 2021. "Chapter 1 - Foundations of demand estimation," The Handbook of Industrial Organization, Editor(s): Kate Ho, Ali Hortaçsu, Alessandro Lizzeri, Elsevier, Volume 4, Issue 1, 2021, 1-62. (Chapter 6)
- Simulation-based estimation: Train's book "Discrete Choice Methods with Simulation" (Chapter 13).

Why Micro Data?

- Micro data links individual characteristics to individual consumer choices
  - individuals with children with Volkswagen Golf
  - rich individuals are less sensitive to price
- With market level data, we learn about the marginal distributions of demographics and choices
- With micro data, we learn about their joint distribution

Why Micro Data?

- Using micro data, we can exploit observable variation between demographics and choices to estimate substitution patterns
- Often consumer/choice-specific variables will involve interactions between consumer attributes and product attributes
- For example, many applications have utilized measures of consumer×product-specific distances (consumer distance to retailer/hospital/etc j affects utility from choosing j but not choosing k)
- Within a market, product×market-level demand shocks are fixed but choices vary with (observable) consumer attributes
  - Note that this "within market variation" has no endogeneity problem.
- For how micro data can exactly help in identification, see Berry and Haile (2021)

Estimation Using Micro Data

Consider the following BLP-stype preferences for consumer i, good j in market m at time t:

$$u_{ijt} = x_{jt}\beta_{it} - \alpha_0 p_{jt} + \xi_{jt} + e_{ijt} \tag{2}$$

- e_{ijmt} type 1 extreme value distributed
- $\beta_{it} = \beta_0 + \Pi d_{it} + \sigma\nu_{it}$ where d_{iy} are demographics (observed at the consumer level in micro data)

Estimation Using Micro Data

Probability that each consumer i choosing j in market m at time t:

$$s_{ijt} = \int \frac{\exp[\delta_{jt} + u_{ijt}(\nu_{it}; \Pi, \sigma)]}{1 + \sum_k \exp[\delta_{kt} + u_{ikt}(\nu_{it}; \Pi, \sigma)]} dF_\nu(\nu_{it})$$

where

- $\delta_{jt} = x_{jt}\beta_0 - \alpha_0 p_{jt} + \xi_{jt}$
- $u_{ijt}(\nu_{it}; \Pi, \sigma) = x_{jt}(\Pi d_{it} + \sigma \nu_{it})$

Estimation Using Micro Data

$$s_{ijt} = \int \frac{\exp[\delta_{jt} + u_{ijt}(\nu_{it}; \Pi, \sigma)]}{1 + \sum_k \exp[\delta_{kt} + u_{ikt}(\nu_{it}; \Pi, \sigma)]} dF_\nu(\nu_{it})$$

- As before, we need to simulate from F_ν
- Unlike before, we observe consumer demographics d_{im} in micro data

Estimation Using Micro Data

One could estimate (δ, Π, σ) by maximizing the product of these (simulated) likelihoods over all consumers

$$L(\delta, \Pi, \sigma) = \prod_{i,t} \int \frac{\exp[\delta_{jt} + u_{ijt}(\nu_{it}; \Pi, \sigma)]}{1 + \sum_k \exp[\delta_{kt} + u_{ikt}(\nu_{it}; \Pi, \sigma)]} dF_\nu(\nu_{it})$$

Estimation Using Micro Data

- To answer most economic questions one will also need to estimate the parameters α_0 and β_0.
- The simplest approach is to run a second-step linear IV regression of the estimated δ_{jt} on x_{jt} and p_{jt}.
- Usual IV's for the price such as cost shifters, BLP instruments and Waldfogel instruments (demographics that are not elements of x_{jt} and not correlated with ξ_{jt})

Estimation Using Micro Data

- When there are numerous products and/or markets, estimation of a large number of constants δ_{jt} can be difficult
- Solutions:
  - δ_{jt} can be also considered as functions of the other parameters, calculated to equate predicted and actual shares at any given values of the other parameters (BLP contraction; see Train's book)

Estimation Using Micro Data

- When there are numerous products and/or markets, estimation of a large number of constants δ_{jt} can be difficult
- Solutions:
  - δ_{jt} can be also considered as functions of the other parameters, calculated to equate predicted and actual shares at any given values of the other parameters (BLP contraction; see Train's book)
  - Treat as parameters with two levels of optimization (Griego et al., 2023): 1.) the inner one: estimate δ_{jt} as a function of other parameters - can be done separately for each market, 2.) the outer one: estimate other parameters.

Zero Market Shares

- Especially with high-frequency micro data (but also more generally), we may observe few, if any, sales for a large number of products, for example in online retail
- Potential solutions:
  - Aggregate data over markets or products until observations with zero sales disappear $\rightarrow$ smooths out heterogeneity, might loose interesting variation
  - Delete zeros $\rightarrow$ creates selection bias, tends to result in estimating consumers as too price inelastic
  - Add a small number of the market share $\rightarrow$ bias to parameter estimates, especially when choice prob. is small

Zero Market Shares

- Quan and Williams (2018) provide a more specialized solution to this problem when products appear in multiple markets
  - Augment the demand model with across-market random effects
  - Decomposes product-level unobservables into a mean national-product level fixed effect and local market deviations
  - The estimator does not rely on the local market shares of individual products, the vast majority of which are zero at each location, but allows to retain information about the distribution of local demand heterogeneity
- See also another solution based on moment inequalities by Gandhi et al. (2023)

Micro Data: Key Takeaways

- If you have micro data, it's almost always a good idea to use them!
- With micro data you need to think about whether/to what extent the individual specific data you have is enough to capture the richness of choices
- Need to think about what type of unobservables you might need to make the model rich enough

Some Useful Literature

- Berry, S., and P. Haile. 2021. "Chapter 1 - Foundations of demand estimation," The Handbook of Industrial Organization, Editor(s): Kate Ho, Ali Hortaçsu, Alessandro Lizzeri, Elsevier, Volume 4, Issue 1, 2021, 1-62.
- Simulation-based estimation: Train's book "Discrete Choice Methods with Simulation" (Chapter 13).
- Gandhi, A., Lu, Z. and Shi, X. (2023), Estimating demand for differentiated products with zeroes in market share data. Quantitative Economics, 14: 381-418.
- Grieco, P., Murry, C., Pinkse, J., Sagl, S. (2023) "Conformant and Efficient Estimation of Discrete Choice Demand Models", Unpublished Manuscript.
- Quan, Thomas W., and Kevin R. Williams (2018) "Product Variety, across-Market Demand Heterogeneity, and the Value of Online Retail." The RAND Journal of Economics, vol. 49, no. 4: 877–913.

Petrin (2002): Combination of Aggregate and Micro Data

Questions

- What are the welfare gains from innovation?
- How much of these gains are captured by consumers and the innovator?
- How big is the extent of first-mover advantage (for the innovator) and the profit cannibalization that took place both initially by the innovator and later by the imitators?

Questions

- What are the welfare gains from innovation?
- How much of these gains are captured by consumers and the innovator?
- How big is the extent of first-mover advantage (for the innovator) and the profit cannibalization that took place both initially by the innovator and later by the imitators?
- The results suggest that the introduction generated large welfare gains for consumers and surplus for the innovator at the expense of the other producers

Setting: Case of the Chrysler's Minivan

Source: Wikipedia

New Alternative for Station Wagons

Source: Wikipedia

Introduction of Minivans

TABLE 3
FAMILY VEHICLE SALES AS A PERCENTAGE OF TOTAL VEHICLE SALES:
U.S. AUTOMOBILE MARKET, 1981–93

Year	Minivans (1)	Station Wagons (2)	Sport-Utilities (3)	Full-Size Vans (4)	Minivans and Station Wagons (5)	U.S. Auto Sales (Millions) (6)
1981	.00	10.51	.58	.82	10.51	7.58
1982	.00	10.27	.79	1.17	10.27	7.05
1983	.00	10.32	3.51	1.04	10.32	8.48
1984	1.58	8.90	5.51	1.20	10.48	10.66
1985	2.32	7.33	6.11	1.05	9.65	11.87
1986	3.63	6.70	5.73	.85	10.43	12.21
1987	4.86	6.47	6.44	.73	11.33	11.21
1988	5.97	5.14	7.18	.69	11.11	11.76
1989	6.45	4.13	7.47	.61	10.58	11.06
1990	7.95	3.59	7.78	.27	11.54	10.51
1991	8.29	3.05	7.80	.29	11.34	9.75
1992	8.77	3.07	9.33	.39	11.84	10.12
1993	9.93	3.02	11.66	.29	12.95	10.71

Micro BLP (Petrin 2002)

Empirical approach

- Random coefficient discrete choice model (BLP)
- Additional micro moments
  - Augment the market-level data on sales and characteristics with information that relates the average demographics of consumers to the characteristics of the products they purchase
  - Information on aggregates of purchasers of new cars

Combining Aggregate and Micro Data

- The extra information plays the same role as consumer-level data, allowing estimated substitution patterns and (thus) welfare to directly reflect demographic-driven differences in tastes for observed characteristics
- For example, observing average family size conditional on the purchase of a minivan and asking the model to reproduce this same average helps to more precisely identify the taste term relating families and minivans
- Similarly, matching probabilities of purchase conditioned on different income levels helps to identify income effects

Combining Aggregate and Micro Data

- Market/product level data (as in BLP):
  - Characteristics and market shares for all new vehicles marketed in the United States from the years 1981 to 1993 (with sales over 1,000 vehicles)
  - Household demographics for representative sample of U.S. population from the Consumer Expenditure Survey (CEX)
- Additional micro data from CEX:
  - Demographics of purchasers of new vehicles linked to the vehicles they purchase for smaller sample (30,000 households)

Utility Specification

The utility specification of consumer i for good j

$$u_{ijt} = \alpha_i \ln(y_i - p_{jt}) + x_{jt}\beta_{it} + \xi_{jt} + e_{ijt} \tag{3}$$

Utility Specification

The utility specification of consumer i for good j

$$u_{ijt} = \alpha_i \ln(y_i - p_{jt}) + x_{jt}\beta_{it} + \xi_{jt} + e_{ijt} \tag{3}$$

$$\alpha_i = \begin{cases} \alpha_1 & \text{if } y_i < \hat{y}_1 \\ \alpha_2 & \text{if } \hat{y}_1 \leq y_i < \hat{y}_2 \\ \alpha_3 & \text{if } y_i \geq \hat{y}_2 \end{cases}$$

where $\hat{y}_1$ and $\hat{y}_2$ divide the U.S. population into three equally sized income groups

Utility Specification

The utility specification of consumer i for good j

$$u_{ijt} = \alpha_i \ln(y_i - p_{jt}) + x_{jt}\beta_{it} + \xi_{jt} + e_{ijt} \tag{3}$$

$$\alpha_i = \begin{cases} \alpha_1 & \text{if } y_i < \hat{y}_1 \\ \alpha_2 & \text{if } \hat{y}_1 \leq y_i < \hat{y}_2 \\ \alpha_3 & \text{if } y_i \geq \hat{y}_2 \end{cases}$$

where $\hat{y}_1$ and $\hat{y}_2$ divide the U.S. population into three equally sized income groups

$$\beta_{it}^k = \begin{cases} \beta^k + \gamma^k \ln(fs_i) v_{it} & k \text{ for minivan, station wagons (+2 other car type) dummies} \\ \beta^k + \gamma^k v_{it} & \text{for all other } k \end{cases} \tag{4}$$

For example, γ^k for the minivan dummy is a taste shifter that allows families of different sizes (fs_i) to value minivans differently

Utility Specification

The utility specification of consumer i for good j

$$u_{ijt} = \alpha_i \ln(y_i - p_{jt}) + x_{jt}\beta_{it} + \xi_{jt} + e_{ijt} \tag{5}$$

$$\alpha_i = \begin{cases} \alpha_1 & \text{if } y_i < \hat{y}_1 \\ \alpha_2 & \text{if } \hat{y}_1 \leq y_i < \hat{y}_2 \\ \alpha_3 & \text{if } y_i \geq \hat{y}_2 \end{cases}$$

$$\beta_{it}^k = \begin{cases} \beta^k + \gamma^k \ln(fs_i) v_{it} & k \text{ for minivan, station wagons (+2 other car type) dummies} \\ \beta^k + \gamma^k v_{it} & \text{for all other } k \end{cases} \tag{6}$$

e_{ijt} type 1 extreme value distributed, v_{it} e.g. normal/log-normal distributed

Moments

- In BLP, we had the following moment condition in the GMM objective:

$$E[\xi_{jt}|z_{jt}] = 0 \rightarrow E[\xi_{jt}' z_{jt}] = 0 \quad (7)$$

- Additional micro moments in GMM to match the average model predictions to the observed averages in the data for the outcomes

Additional Micro Moments

- Petrin also considers the following moments:
  1. Probability of purchasing a new vehicle, given income (discretized into three bins)

$$E(\text{purchasing a new vehicle}|y_i < \hat{y}_1)$$

$$E(\text{purchasing a new vehicle}|\hat{y}_1 \leq y_i < \hat{y}_2)$$

$$E(\text{purchasing a new vehicle}|y_i \geq \hat{y}_2)$$

  2. Average family size (fs_i) given the type of vehicle purchased

$$E(fs_i|\text{purchase } j), j = \text{minivan, station wagon, etc}$$

  3. Probability the head of household is 30–60 years old, given the type of vehicle purchased

Additional Micro Moments

- How would you predict e.g. $E(fs_i|\text{purchase } j), j =$minivan, station wagon, etc?

Additional Micro Moments

- How would you predict e.g. $E(fs_i|\text{purchase } j), j =$minivan, station wagon, etc?
- Denote "i purchases j in market t" by $a_{it} = j$
- More specifically, $E(fs_i|a_{it} = j) \equiv E(fs_i|a_{it} = j, \delta_t, \theta_2)$
- Bayes rule: $E(fs_i|a_{it} = j, \delta_t, \theta_2) = \int fs_i dF(fs_i|a_{it} = j, \delta_t, \theta_2) = \int fs_i \frac{P(a_{it}=j|fs_i,\delta_t,\theta_2)}{P(a_{it}=j|\delta_t,\theta_2)} dF(fs_i)$
- For estimation, just stack the micro moments (matching model predictions to data) with the original IV moments ( ) and run BLP as usual
- Also have a supply side model of oligopolistic competition (pricing decisions, no formal model of entry/exit/innovation)

Demand Estimates

TABLE 4
PARAMETER ESTIMATES FOR THE DEMAND-SIDE EQUATION

Variable	OLS Logit (1)	Instrumental Variable Logit (2)	Random Coefficients (3)	Random Coefficients and Microdata (4)
	A. Price Coefficients (α's)			
α_1	.07	.13	4.92	7.52
	(.01)**	(.01)**	(9.78)	(1.24)**
α_2			11.89	31.13
			(21.41)	(4.07)**
α_3			37.92	34.49
			(18.64)**	(2.56)**

Demand Estimates

TABLE 5
RANDOM COEFFICIENT PARAMETER ESTIMATES

	Random Coefficients (γ's)	
Variable	Uses No Microdata (1)	Uses CEX Microdata (2)
Constant	1.46	3.23
	(.87)*	(.72)**
Horsepower/weight	.10	4.43
	(14.15)	(1.60)**
Size	.14	.46
	(8.60)	(1.07)
Air conditioning standard	.95	.01
	(.55)*	(.78)
Miles/dollar	.04	2.58
	(1.22)	(.14)**
Front wheel drive	1.61	4.42
	(.78)**	(.79)**
γ_{mi}	.97	.57
	(2.62)	(.10)**
γ_{sw}	3.43	.28
	(5.39)	(.09)**
γ_{su}	.59	.31
	(2.84)	(.09)**
γ_{pv}	4.24	.42
	(32.23)	(.21)**

NOTE.—The OLS and instrumental variable models assume that these random coefficients are zero. Standard errors are in parentheses. A quadratic time trend is included in all specifications. The subscript *mi* stands for minivan, *sw* for station wagon, *su* for sport-utility, and *pv* for full-size passenger van.
* Z-statistic >1.
** Z-statistic >2.

Counterfactuals

- Benefits or harms of the minivan introduction

Counterfactuals

- Benefits or harms of the minivan introduction
  - to consumers
  - to the innovator (Chrysler)
  - to competitors (e.g. General Motors with station wagon)
  - to imitating firms (almost all other)

Counterfactuals

- Measure changes in welfare (consumer welfare, profits) from the minivan introduction

Counterfactuals

- Measure changes in welfare (consumer welfare, profits) from the minivan introduction
  - Take minivan out of the market
  - Simulate new market shares, profits, and welfare
  - For consumers:
    - draw D_i, v_i, e_{ij}
    - calculate choice and utility with full choice set (with minivans) and with reduced choice set (without minivans)

Counterfactuals

- Measure changes in welfare (consumer welfare, profits) from the minivan introduction
  - Take minivan out of the market
  - Simulate new market shares, profits, and welfare
  - For consumers:
    - draw D_i, v_i, e_{ij}
    - calculate choice and utility with full choice set (with minivans) and with reduced choice set (without minivans)
    - calculate compensating variation CV_i as a measure changes in consumer welfare from the introduction of the minivan
    - CV_i=dollar amount the consumer would need to be just indifferent between the equilibrium with minivans and the one without them

Profits

TABLE 11
CHANGE IN INDUSTRY AND BIG THREE TOTAL VARIABLE PROFITS WITH THE ADVENT OF MINIVANS

YEAR	INDUSTRY	CHRYSLER		FORD		GM	
1984	−.21%	\$202.5	14.38%	−\$31.8	−1.16%	−\$155.8	−1.50%
1985	−.13%	\$259.1	13.99%	−\$37.4	−1.29%	−\$171.0	−1.63%
1986	.14%	\$201.1	12.42%	\$54.7	1.84%	−\$119.9	−1.09%
1987	.17%	\$346.1	23.27%	−\$22.8	−.66%	−\$174.5	−2.14%
1988	.65%	\$504.1	32.50%	−\$24.7	−.70%	−\$235.4	−2.90%

NOTE.—Dollar figures are given in millions. The numbers are computed using the model to estimate profits both with minivans in the market and with minivans removed from the market (see Sec. V).

TABLE 12
CHRYSLER'S PROFIT DISSIPATION WITH ENTRY OF FORD AND GM MINIVANS

YEAR	CHANGE IN TOTAL VARIABLE PROFITS	
1985	−\$6.06	−.16%
1986	−\$22.72	−1.99%
1987	−\$42.35	−2.25%
1988	−\$55.68	−2.63%

NOTE.—These profit changes are computed using the model (see the text).

Welfare

TABLE 13

CHANGE IN U.S. WELFARE FROM THE MINIVAN INNOVATION, 1984–88 ($ Millions)

Year	Compensating Variation	Change in Producer Profits	Welfare Change
1984	367.29	−36.68	330.61
1985	625.04	−25.07	599.97
1986	439.93	27.30	467.23
1987	596.59	29.75	626.34
1988	775.70	110.24	885.94
Total	2,804.55	105.54	2,910.09

NOTE.—Computations were done using 1982–84 CPI-adjusted dollars.

Part 1+2: Summary

- BLP give us both a statistical estimator and an algorithm to obtain estimates.
- Attractive for many differentiated products markets
- Flexible substitution patterns, addressing endogeneity concerns
- Computationally demanding, progress on these challenges
  - Alternative algorithms (MPEC), improvements in the BLP NFP algorithm
  - Current best practice: pyBLP

Part 1+2: Summary

- Many variations of the model in the literature, e.g. using different data structures
  - Aggregate data (original BLP)
  - Micro data
  - Panel data
  - Hybrid
- Often a good idea to use micro data if you have them
- Help in identification, provide additional variation to pin down substitution patterns/heterogeneity in consumer preferences
- Later: identification, more about IVs, supply side

Any Questions?

Email: tanja.saxell@aalto.fi