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

Lecture 3

Tanja Saxell

Practicalities

- Exercises, slides, and reading liston MyCourses (https://mycourses.aalto.fi/course/view.php?id=39506)
- Return your exercise answers to MyCourses as a pdf or html and include the code
- Exercise nbr 1 (logit, nested logit), due 1 Feb
- Questions about practicalities or exercises? Email to Helena Rantakaulio (helena.rantakaulio@aalto.fi)
- My email: tanja.saxell@aalto.fi

Outline: Previous Lectures (Toivanen)

- Discrete choice models
  - Logit
  - Nested logit
- These models are intuitive and easy to implement.
- Why not enough?

Outline: Lectures 3-4 (Saxell)

- Reminder: problems with logit/nested logit
- Random coefficients logit - consumer heterogeneity in preferences, rich substitution patterns
  - Estimator
  - Algorithm
  - Aggregate product-level data
- Challenges and extensions such as
  - Micro data
  - Combination of aggregate and micro data

Literature: Lectures 3-4 (Saxell)

- Key papers:
  - Aggregate data: BLP = Berry, Steven, Levinsohn, J., and Pakes, A. (1995). "Automobile Prices in Market Equilibrium," Econometrica, 63(4), 841–890.
  - Aggregate + 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.
- For practitioners:
  - Nevo, Aviv. 2000. "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," Journal of Economics & Management Strategy, 9(4), 513–548.
  - Conlon, C. and Gortmaker, J. 2020. Best practices for differentiated products demand estimation with PyBLP. The RAND Journal of Economics, 51: 1108-1161.

Outline: Later Lectures (5-8)

- Lectures 5-6 (Toivanen): Supply side
- Lectures 7-8 (Vehviläinen): Mergers and market power

Problems with Logit

- Independence of irrelevant alternatives (IIA)
  - IIA states that the probability of choosing one product over another does not depend on the presence or absence of other "irrelevant" alternatives.
  - In other words, whatever else is on offer does not matter in the choice between j and m.
  - Implication: when the probability of choosing a given alternative changes, all other choice probabilities change in proportion.

Problems with Logit

- Often unrealistic price elasticities:

$$e_{jk} = \begin{cases} -\alpha p_j(1 - s_j) & \text{if } k = j \\ \alpha p_k s_k & \text{if } k \neq j. \end{cases} \tag{1}$$

- The own price elasticity e_{jj} is increasing in price (absolute value). Why this is unrealistic - given an example!

Problems with Logit

- Often unrealistic price elasticities:

$$e_{jk} = \begin{cases} -\alpha p_j(1 - s_j) & \text{if } k = j \\ \alpha p_k s_k & \text{if } k \neq j. \end{cases} \tag{1}$$

- The own price elasticity e_{jj} is increasing in price (absolute value). Why this is unrealistic - given an example!
  - We would think people who buy expensive products are less sensitive to price.
- The cross-elasticity e_{jk}, $k \neq j$ depends only on market share and price of k (but not j!) but not on similarities between goods (IIA).

Source of the Problem

- Source of the problem: no correlation in the preference shock across products.
    - E.g., when the preference shock to BMW is high, the preference shock to Mercedes Benz should also be high, while the preference shock to Fiat should be relatively independent.
- Therefore, the preference shocks between two alternatives should be more correlated when they are closer in the characteristics space.
- Most of the extensions try to correct for the above.

Why Important?

The main reason to estimate demand is to quantify demand parameters/elasticities:

- determine responses to and welfare effects of price changes
- determine responses to counterfactual policies (mergers, entry, tax changes etc.)
- used with a supply model to infer markups and market power

Solutions

- Nested logit: assume a particular correlation structure among the structural errors e_{ij}. Within a nest, alternatives are "closer substitutes" than across-nest alternatives.
- Extensions: multi-level tree structure.
- One big problem with nested-logit: need to a-priori group products to nests, this is not trivial (examples?).

Solutions

- Nested logit: assume a particular correlation structure among the structural errors e_{ij}. Within a nest, alternatives are "closer substitutes" than across-nest alternatives.
- Extensions: multi-level tree structure.
- One big problem with nested-logit: need to a-priori group products to nests, this is not trivial (examples?).
  - E.g. housing choice: Level 1: Location (Neighborhood), level 2: Housing Type (Rent, Buy, House, Apt); and Level 3: Housing (Bedrooms)?
  - Or some other combination of these?
- Different nest structures can produce very different results.
- The random coefficients models will try to solve this and provide more general treatment.

Random coefficients model: Berry, Levisohn and Pakes or BLP (1995)

BLP

- Workhorse empirical model of demand (and supply) of differentiated products.
- Many of the ideas in Berry (1994), mostly for simpler models (without random coefs.).
- Many extensions and variations.

BLP

- Random coefficients with individuals heterogeneity $\rightarrow$ rich substitution patterns.
- Requires only aggregate (product and market) level data.
  - Because we can construct the aggregated data from individual level data, all the arguments should go through with the individual choice level data.

BLP

- Explicit about unobservables (to the econometrician), including the nature of endogeneity problem.
  - For example, the econometrician may not observe brand values that are created by advertisement and perceived by consumers.
  - Such unobserved product characteristics are likely to be correlated with the price.
- Use the model to reveal appropriate instruments (based on market competition).
- Propose an algorithm for consistent estimation of the model and standard errors.

Demand Model

- A consumer chooses one of the available options (unit demand).
- There are J differentiated products or inside goods (e.g., different types of cars) $j = 1, .., J$.
- One options should be the outside good, ($j = 0$) i.e. none of the products above (e.g., do not buy a car).
- Note that the model is fairly general, a single option could also be a product bundle, e.g. milk+cheese, shirt+jeans...

Demand Model

- Specification for the conditional indirect utility of consumer i for product j in market t:

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

  - x_{jt}, p_{jt}: observable product/market characteristics
  - ξ_{jt} unobserved product/market characteristic (demand shock e.g. brand/quality, structural errors on the demand side).
  - e_{ijt} idiosyncratic taste for the product.
- Only utility differences matter, so need a normalization for the outside good, e.g. $u_{i0t} = e_{i0t}$.

Outside Good and Income Effect

- Could also have the utilities to depend on income y_i
- How the preference for the outside good is modeled determines how the individual income affects the choice
- For example, assume (Nevo, 2000)

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

$$u_{i0t} = \alpha y_i + \xi_{0t} + e_{ijt} \tag{4}$$

- The income level does not affect the choice because the term is common and constant across choices (there is no income effect)
- This is in contrast to the case where we have $\alpha \ln(y_i - p_j)$ instead of $\alpha(y_i - p_j)$ in ( ) and $\alpha \ln(y_i)$ instead of αy_i in ( ) as in BLP
- For simplicity, assume that there is no income effect

Heterogeneity in Preferences

Utility specification: $u_{ijt} = x_{jt}\beta_{it} - \alpha p_{jt} + \xi_{jt} + e_{ijt}$

- Random coefficients: $\beta_{it} = \beta + \sigma v_{it}$
  - For simplicity, just for x_{jt} but could be also for p_{jt}
- e_{ijt}, v_{it} iid across consumers and markets, often:
  - e_{ijt}: iid type 1 extreme value distributed (logit)
  - v_{it}: $N(0, 1)$ or drawn from the distribution of demographics (e.g. income) in market t (mean and std observed in aggregate data)

Random Coefficients

- Products differ in different ways, consumers have heterogeneous preferences over these differences.
- For example, consumers with strong taste for one electricity car will probably like other electricity cars too.
- Random coefficients on product characteristics can capture this.
- Large β_{it}^{k}, strong taste for characteristic x^k
- Consumer i's first (and also second) choice have high values of x^k.
- Key issue as a reminder: produces more sensible substitution patterns.
    - As a result, the degree of competition depends on the degree to which similar products are available.

Data (BLP)

- The data set includes information on (essentially) all car models marketed during the 20 year period beginning in 1971 and ending in 1990.
- Unbalanced panel: car models both appear and exit over this period.
- Identify retail list prices (transaction prices are not easy to find) and other product characteristics.
- Distinguishes which firms produce which model
    - Crucial for the supply model and also for the IVs
- In total, N=2217 model/year observations.

Available Products (BLP)

TABLE 1

DESCRIPTIVE STATISTICS

Year	No. of Models	Quantity	Price	Domestic	Japan	European	HP/Wt	Size	Air	MPG	MP$
1971	92	86.892	7.868	0.866	0.057	0.077	0.490	1.496	0.000	1.662	1.850
1972	89	91.763	7.979	0.892	0.042	0.066	0.391	1.510	0.014	1.619	1.875
1973	86	92.785	7.535	0.932	0.040	0.028	0.364	1.529	0.022	1.589	1.819
1974	72	105.119	7.506	0.887	0.050	0.064	0.347	1.510	0.026	1.568	1.453
1975	93	84.775	7.821	0.853	0.083	0.064	0.337	1.479	0.054	1.584	1.503
1976	99	93.382	7.787	0.876	0.081	0.043	0.338	1.508	0.059	1.759	1.696
1977	95	97.727	7.651	0.837	0.112	0.051	0.340	1.467	0.032	1.947	1.835
1978	95	99.444	7.645	0.855	0.107	0.039	0.346	1.405	0.034	1.982	1.929
1979	102	82.742	7.599	0.803	0.158	0.038	0.348	1.343	0.047	2.061	1.657
1980	103	71.567	7.718	0.773	0.191	0.036	0.350	1.296	0.078	2.215	1.466
1981	116	62.030	8.349	0.741	0.213	0.046	0.349	1.286	0.094	2.363	1.559
1982	110	61.893	8.831	0.714	0.235	0.051	0.347	1.277	0.134	2.440	1.817
1983	115	67.878	8.821	0.734	0.215	0.051	0.351	1.276	0.126	2.601	2.087
1984	113	85.933	8.870	0.783	0.179	0.038	0.361	1.293	0.129	2.469	2.117
1985	136	78.143	8.938	0.761	0.191	0.048	0.372	1.265	0.140	2.261	2.024
1986	130	83.756	9.382	0.733	0.216	0.050	0.379	1.249	0.176	2.416	2.856
1987	143	67.667	9.965	0.702	0.245	0.052	0.395	1.246	0.229	2.327	2.789
1988	150	67.078	10.069	0.717	0.237	0.045	0.396	1.251	0.237	2.334	2.919
1989	147	62.914	10.321	0.690	0.261	0.049	0.406	1.259	0.289	2.310	2.806
1990	131	66.377	10.337	0.682	0.276	0.043	0.419	1.270	0.308	2.270	2.852
All	2217	78.804	8.604	0.790	0.161	0.049	0.372	1.357	0.116	2.099	2.086

Note: The entry in each cell of the last nine columns is the sales weighted mean.

Substitution to Outside Good (BLP)

TABLE VII

SUBSTITUTION TO THE OUTSIDE GOOD

Model	Given a price increase, the percentage who substitute to the outside good (as a percentage of all who substitute away.) Logit	BLP
Mazda 323	90.870	27.123
Nissan Sentra	90.843	26.133
Ford Escort	90.592	27.996
Chevy Cavalier	90.585	26.389
Honda Accord	90.458	21.839
Ford Taurus	90.566	25.214
Buick Century	90.777	25.402
Nissan Maxima	90.790	21.738
Acura Legend	90.838	20.786
Lincoln Town Car	90.739	20.309
Cadillac Seville	90.860	16.734
Lexus LS400	90.851	10.090
BMW 735*i*	90.883	10.101

Exogenous and Endogenous Characteristics

Utility specification: $u_{ijt} = x_{jt}\beta_{it} - \alpha p_{jt} + \xi_{jt} + e_{ijt}$

- exogenous product characteristics x_{jt} (uncorrelated with ξ_{jt})
- endogenous product characteristics p_{jt}, usually the price
  - firms know ξ_{jt} when setting prices.
  - each price depends on $\xi_t = (\xi_{1t}, ..., \xi_{Jt})$.
  - need instruments
- But note that we dot estimate the equation above, utilities are not observable.
- Observed prices and quantities/market shares are both endogenous (simultaneously determined).

Utility Specification: Mean Utility

- Redefine the utility specification:

$$
\begin{aligned}
u_{ijt} &= x_{jt}\beta_{it} - \alpha p_{jt} + \xi_{jt} + e_{ijt} && (5) \\
&= \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}, \theta_1) + u_{ijt}(x_{jt}, v_{it}, \theta_2) + e_{ijt}, && (6)
\end{aligned}
$$

where

- $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ is the mean utility of product j in market t, with $\delta_{0t} = 0$ (normalization)
- a deviation from that mean:

$$
\begin{aligned}
\mu_{ijt} &= u_{ijt} + e_{ijt} && (7) \\
&= x_{jt}\sigma v_{it} + e_{ijt} && (8) \\
&= x_{jt}\tilde{\beta}_i + e_{ijt} && (9)
\end{aligned}
$$

- $\theta_1 = (\beta, \alpha)$, $\theta_2 = \sigma$

Consumer Choice and Market Share

- Consumer i''s choice:

$$a_{it} = \arg\max_j u_{ijt}.$$

.

Consumer Choice and Market Share

- Consumer i''s choice:

$$a_{it} = \arg\max_j u_{ijt}.$$

.

- The market share of product j is just an integral over the mass of consumers in the region A_{jt}:

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

where

$$A_{jt}(\delta_t, x_t, \theta_2) = \{(v_{it}, e_{i0t}, ..., e_{iJt}) : u_{ijt} \geq u_{ikt} \text{ for all } k \in \{0, ..., J\}\}$$

Consumer Choice and Market Share

With Type 1 extreme value distributed error terms (e) and random coefficients, the predicted market share is:

$$s_j(\delta_t, x_t, \theta_2) = \int \frac{\exp[\delta_{jt} + x_{jt}\tilde{\beta}_i]}{1 + \sum_k \exp[\delta_{kt} + x_{kt}\tilde{\beta}_i]} dF_{\tilde{\beta}}(\tilde{\beta}_i|\theta_2)$$

Total Demand for Each Product

- If M_t is a measure of the total number of potential consumers in market t, the total demand for product j is in market t:

$$q_{jt} = M_t \times s_j(\delta_t, x_t, \theta_2) \tag{10}$$

- And for the outside good:

$$q_{0t} = M_t - \sum_{j=1}^{J} M_t \times s_j(\delta_t, x_t, \theta_2) \tag{11}$$

Data and Estimation

Data

- Market and product level data (observable): x_t, p_t, s_t, M_t and z_t (instruments).
  - Could also use aggregate data on demographics such as income (later).
- How would you measure M_t and the market share of the outside good $s_{0t} = 1 - \sum_{j=1}^{J} s_{jt}/M_t$?

Data

- Market and product level data (observable): x_t, p_t, s_t, M_t and z_t (instruments).
  - Could also use aggregate data on demographics such as income (later).
- How would you measure M_t and the market share of the outside good $s_{0t} = 1 - \sum_{j=1}^{J} s_{jt}/M_t$?
  - BLP: M_t is the number of households in the U.S. taken for each year from the Statistical Abstract of the U.S.
- Perform robustness checks on market size assumptions; might matter a lot for the estimates and outcomes!
  - See a job market paper by Zhang, L. (2023): "Identification and Estimation of Market Size in Discrete Choice Demand Models."

Estimation

- Assume that θ_2 (and the distributions F_v and F_e) are already known.
- For each market t, find $\delta_t \in R^J$ such that $s_j(\delta_t, x_t, \theta_2) = s_{jt}\ \forall j$.
  - Invert market shares to recover mean utilities δ_t.
  - Done this way, δ_t is such that the predicted market shares fit the observed market shares exactly.

Estimation: To-Do

- Instruments
- Inversion step: from market shares to mean utilities
- Formally, define an estimator and algorithm
- (Add supply model)

Price Endogeneity

- Identification concerns: price endogeneity (correlates with ξ_{jt}).
- Need instruments: variables that exogenously shift prices and quantities independently.

BLP Instruments

- IVs based on market competition.
    - In oligopolistic competition, firm j sets the price as a function of characteristics of products produced by competing firms.
    - However, characteristics of competing products should not depend on a consumer's valuation of firm j's product.
    - Similarly, for multiproduct firms, can construct IVs using characteristics of all other products produced by same firm j.

BLP Instruments

- The following are used as IVs for the price of product in a given market, p_{jt}

$$\sum_{k \neq j \in \mathcal{J}_t \cap \mathcal{F}_f} x_{kt},$$

$$\sum_{k \neq j \in \mathcal{J}_t \setminus \mathcal{F}_f} x_{kt}.$$

  where f is the firm that owns product j and $\mathcal{F}_f$ is the set of products firm f owns
- For example, if one of the characteristics is the size of a car, then the IVs for product j includes the sum of size across own-firm products and the sum of size across rival firm products

Alternative Instruments

- Traditional cost shifters; however need variation in costs across alternatives
- Proxies of cost shifters; price of the same product in other markets (Hausman instruments), valid if demand shocks are uncorrelated across markets
- Characteristics of nearby markets (Waldfogel instruments, after Waldfogel 2003)
- Exogenous shifters of market structure (e.g., firm ownership) that affect prices through equilibrium markups
- More on instruments and identification later (Hyytinen, lectures 5-6)

Inversion Step

- Find δ_t solves the nonlinear system $s_t = s(\delta_t, x_t, \theta_2)$, or equivalently

$$\delta_t = \delta_t + \underbrace{ln(s_t)}_{\textit{Data!}} - \underbrace{ln(s(\delta_t, x_t, \theta_2))}_{\textit{Model prediction!}} \tag{12}$$

- They show that under mild conditions on the linear random coef. random utility model, $T(\delta_t) = \delta_t + ln(s_t) - ln(s(\delta_t, x_t, \theta_2))$ is a contraction mapping.
- This means that
  - it has a (unique) fixed point in δ_t.
    - $s_t = s(\delta_t, x_t, \theta_2)$ has an inverse $\delta_t = D^{-1}(s_t, x_t, \theta_2)$.
    - We can therefore perform a non-linear change of variables from observed market shares (s_t), x_t and θ_2 to δ_t (see Berry and Haile, 2014).

Analytical Inversion: Logit

Recall the utility specification:

$$u_{ijt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \varepsilon_{ijt}, \quad s_{jt} = \frac{\exp[x_{jt}\beta - \alpha p_{jt} + \xi_{jt}]}{1 + \sum_k \exp[x_{kt}\beta - \alpha p_{kt} + \xi_{kt}]}$$

- ξ_{jt} potentially correlated with price $\text{Corr}(\xi_{jt}, p_{jt}) \neq 0$
- But not characteristics $E[\xi_{jt}|x_{jt}] = 0$.

Analytical Inversion: Logit

Taking logs:

$$\ln(s_{0t}) = -\ln\left(1 + \sum_k \exp[x_{kt}\beta + \xi_{kt}]\right)$$

$$\ln(s_{jt}) = [x_{jt}\beta - \alpha p_{jt} + \xi_{jt}] - \ln\left(1 + \sum_k \exp[x_{kt}\beta + \xi_{kt}]\right)$$

$$\underbrace{\ln(s_{jt}) - \ln(s_{0t})}_{\textit{Data}!} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

Exploit the fact that we have one ξ_{jt} for every share s_{jt} (one to one mapping)

IV Logit Estimation

1. Transform the data: $\ln(s_{jt}) - \ln(s_{0t})$.
2. IV Regression of: $\ln(s_{jt}) - \ln(s_{0t})$ on $x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ with IV z_{jt}.

Analytical Inversion: Nested Logit (Berry 1994)

For nested logit, the same as logit plus an extra term $\ln(s_{j|g})$ the within group share:

$$\underbrace{\ln(s_{jt}) - \ln(s_{0t}) - \sigma \ln(s_{j|gt})}_{\text{Data!}} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

$$\ln(s_{jt}) - \ln(s_{0t}) = x_{jt}\beta - \alpha p_{jt} + \sigma \ln(s_{j|gt}) + \xi_{jt}$$

- Note that $\ln(s_{j|g})$ is also endogenous – we are regressing Y on a function of Y.
- A common instrument is the number of products within the nest.

Inversion: BLP (Random Coefficients)

We can't solve for δ_{jt} analytically this time.

$$s_j(\delta_t, x_t, \theta_2) = \int \frac{\exp[\delta_{jt} + x_{jt}\tilde{\beta}_i]}{1 + \sum_k \exp[\delta_{kt} + x_{kt}\tilde{\beta}_i]} dF_{\tilde{\beta}}(\tilde{\beta}_i|\theta_2)$$

- This is a $J \times J$ system of equations for each t.
- Model predictions $s_j(\delta_t, x_t, \theta_2)$ involve high-dimensional integrals, use simulation (Monte Carlo Integration) to approximate it ("method of simulated moments" instead of GMM).
- There is a unique vector $\boldsymbol{\delta_t}$ that solves it for each market t.
- We can solve δ_t recursively (because the contraction mapping has a unique fixed point) at each trial value of θ_2 (BLP "nested fixed point algorithm").

BLP Estimator (Without Supply Side)

- GMM estimator of $\theta = (\theta_1, \theta_2)$:

$$\min_{\theta} g(\xi(\theta))' W g(\xi(\theta)) \text{ s.t.}$$

  - $g(\xi(\theta)) = \frac{1}{N} \sum_{j,t} \xi_{jt}(\theta)' z_{jt}$
  - $\xi_{jt}(\theta) = \delta_{jt}(\theta_2) - x_{jt}\beta - \alpha p_{jt}$ where $\delta_{jt}(\theta_2) \equiv \delta_j(s_t, x_t, \theta_2)$
  - $s_{jt} = s_j(\delta_t, x_t, \theta_2)$
  - $s_j(\delta_t, x_t, \theta_2) = \int \frac{\exp[\delta_{jt}(\theta_2) + x_{jt}\tilde{\beta}_i]}{1 + \sum_k \exp[\delta_{kt}(\theta_2) + x_{kt}\tilde{\beta}_i]} dF_{\tilde{\beta}}(\tilde{\beta}_i | \theta_2)$, approximation via simulation
  - W: standard GMM weighting matrix: a consistent estimate of $E(z'\xi\xi'z)^{-1}$

BLP Estimator (Without Supply Side)

- GMM estimator of $\theta = (\theta_1, \theta_2)$:

$$\min_{\theta} g(\xi(\theta))' W g(\xi(\theta)) \text{ s.t.}$$

- $g(\xi(\theta)) = \frac{1}{N} \sum_{j,t} \xi_{jt}(\theta)' z_{jt}$
- $\xi_{jt}(\theta) = \delta_{jt}(\theta_2) - x_{jt}\beta - \alpha p_{jt}$ where $\delta_{jt}(\theta_2) \equiv \delta_j(s_t, x_t, \theta_2)$
- $s_{jt} = s_j(\delta_t, x_t, \theta_2)$
- $s_j(\delta_t, x_t, \theta_2) = \int \frac{\exp[\delta_{jt}(\theta_2) + x_{jt}\tilde{\beta}_i]}{1 + \sum_k \exp[\delta_{kt}(\theta_2) + x_{kt}\tilde{\beta}_i]} dF_{\tilde{\beta}}(\tilde{\beta}_i | \theta_2)$, approximation via simulation
- W: standard GMM weighting matrix: a consistent estimate of $E(z'\xi\xi' z)^{-1}$
  - At the true parameter value, θ^*, the moment condition $E(z_t \xi_t(\theta^*)) = 0$
  - The weight matrix defines the metric by which we measure how close to zero we are
  - By using the inverse of the variance-covariance matrix of the moments, we give less weight to those moments that have a higher variance

Contraction: BLP

BLP propose an algorithm to find δ_{jt} s.t. $s_{jt} = s_j(\delta_t, x_t, \theta_2)$. Fix θ_2 and solve for δ_t.

$$\delta_{jt}^{(k)} = \delta_{jt}^{(k-1)} + \left[\underbrace{\ln(s_{jt})}_{\textit{Data}!} - \underbrace{\ln(s_j(\delta_t^{(k-1)}, x_t, \theta_2))}_{\textit{Model prediction!}} \right]$$

- Idea: begin by evaluating the right-hand side of eq. at some initial guess for vector δ_t^0, obtain a new δ_t^1 as the output of this calculation for all j in market t, substitute it back into the right hand side of eq., and repeat this process until convergence.
- If iterate until $|\delta_t^{(k)} - \delta_t^{(k-1)}| < \epsilon_{tol}$ you can recover the δ's so that the observed shares and the predicted shares are identical.
- ϵ_{tol} has to be small (loose tolerance value can make performance poor).
- $s(\delta_t^{(k-1)}, \theta_2)$ requires computing the numerical integral each time (e.g., via monte carlo, later on this).

BLP Algorithm: Basic Idea

- Outer loop: search over trial values of the parameter vector $\theta = (\theta_1, \theta_2)$
- Inner loop: given θ, find a solution for $\delta_t(\theta_2)$ 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 Pseudocode

From the outside, in:

- Outer loop: search over parameters $\theta = (\theta_1, \theta_2)$ to minimize GMM objective:

$$\widehat{\theta_{BLP}} = \arg\min_{\theta} g(\xi(\theta))' W g(\xi(\theta)) \tag{13}$$

BLP Pseudocode

From the outside, in:

- Outer loop: search over parameters $\theta = (\theta_1, \theta_2)$ to minimize GMM objective:

$$\widehat{\theta_{BLP}} = \arg\min_{\theta} g(\xi(\theta))' W g(\xi(\theta)) \tag{13}$$

- Inner Loop:
    - Fix a guess of θ_2.
    - Solve for δ_{jt} which satisfies $s_j(\delta_t, x_t, \theta_2) = s_{jt}$.
        - Simulated moments: computing $s_j(\delta_t, x_t, \theta_2)$ requires numerical integration (simulation).

BLP Pseudocode

From the outside, in:

- Outer loop: search over parameters $\theta = (\theta_1, \theta_2)$ to minimize GMM objective:

$$\widehat{\theta_{BLP}} = \arg\min_{\theta} g(\xi(\theta))' W g(\xi(\theta)) \tag{13}$$

- Inner Loop:
  - Fix a guess of θ_2.
  - Solve for δ_{jt} which satisfies $s_j(\delta_t, x_t, \theta_2) = s_{jt}$.
    - Simulated moments: computing $s_j(\delta_t, x_t, \theta_2)$ requires numerical integration (simulation).
- We can do IV-GMM to recover θ

$$\delta_{jt}(\theta_2) = x_{jt}\beta - \alpha p_{jt} + \xi_{jt} \rightarrow \xi_{jt}(\theta_1, \theta_2)$$

- Use $\hat{\xi}(\theta)$ to approximate $g(\xi(\theta)) \approx \frac{1}{JT} \sum_{j,t} Z_{jt}' \xi_{jt}$
- Plug into the GMM objective function and approximate W ( )
- Iterate until convergence
- Standard errors: standard MSM (method of simulated moments)

Linear and Nonlinear Parameters

- Important simplification: θ_1 enter objective function and $\xi_{jt} \equiv \delta_{jt}(\theta_2) - (x_{jt}\beta - \alpha p_{jt})$ linearly
- Given θ_2 and W, we have closed-form expression for optimal θ_1 (as a function of θ_2)
- Outer loop (nonlinear) search only involves θ_2
    - The nonlinear parameters θ_2 solve for the mean utility levels δ_{jt} that set the predicted market shares equal to the observed market shares

Approximating Market Shares: Numerical Integration

- Model predictions $s_j(\delta_t, x_t, \theta_2)$ involve high-dimensional integrals, a common approach is to use Monte Carlo Integration to approximate them
- MC integration is a technique for numerical integration using random numbers
- Particularly useful for higher-dimensional integrals

Numerical Integration: Example

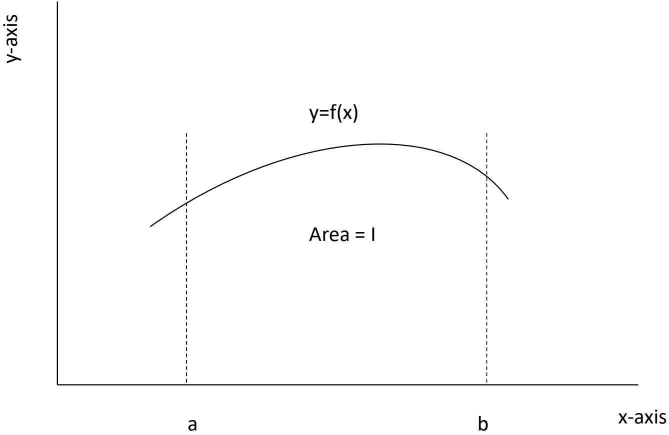

Figure: Approximate $I = \int_a^b f(x)dx$

Numerical Integration: Approximate $I = \int_a^b f(x)dx$

- In the simplest (deterministic) approach, the integral is approximated by a summation over N points at a regular interval $\triangle x$ for x:

$$\hat{I} = \sum_{i=1}^{N} f(x_i) \triangle x$$

where $x_i = a + (i - 0.5) \triangle x$ and $\triangle x = \frac{b-a}{N}$, i.e.

$$\hat{I} = \frac{b-a}{N} \sum_{i=1}^{N} f(x_i)$$

- Takes the value of f from the midpoint of each interval

Numerical Integration: Approximate $I = \int_a^b f(x)dx$

- In the simplest (deterministic) approach, the integral is approximated by a summation over N points at a regular interval $\triangle x$ for x:

$$\hat{I} = \sum_{i=1}^{N} f(x_i) \triangle x$$

where $x_i = a + (i - 0.5) \triangle x$ and $\triangle x = \frac{b-a}{N}$, i.e.

$$\hat{I} = \frac{b-a}{N} \sum_{i=1}^{N} f(x_i)$$

- Takes the value of f from the midpoint of each interval
- The sampling method for MC integration is very similar to the simple approach
- Instead of sampling at regular intervals $\triangle x$, we now sample at random points x_i, and then take the average over NS values of these

BLP: Approximating Market Shares

- Approximation of predicted shares, given θ
  $s_j(\delta_t, x_t, \theta_2) = \int \frac{\exp[\delta_{jt}(\theta_2)+x_{jt}\tilde{\beta}_i]}{1+\sum_k \exp[\delta_{kt}(\theta_2)+x_{kt}\tilde{\beta}_i]} dF_{\tilde{\beta}}(\tilde{\beta}_i|\theta_2)$.
- Draw NS values of v_{it} e.g. from $N(0,1)$ to get $\tilde{\beta}_{it} = \sigma v_{it}$.
- Approximate: $s_j(\delta_t, x_t, \theta_2) \approx \frac{1}{NS} \sum_{i=1}^{NS} \frac{\exp[\delta_{jt}(\theta_2)+x_{jt}\tilde{\beta}_{it}]}{1+\sum_k \exp[\delta_{kt}(\theta_2)+x_{kt}\tilde{\beta}_{it}]}$.
- Use the same set of draws for each value θ (outer loop).

Illustrating Benefits of BLP - Reminder of Problems with Logit

- Logit model had the problem of IIA (independence of irrelevant alternatives)
- Under the IIA, the ratio of choice probabilities between two alternatives depend only on the mean utility of these two alternatives and are independent of irrelevant alternatives

$$\frac{s_j(\delta_t)}{s_l(\delta_t)} = \frac{\exp[x_{jt}\beta - \alpha p_{jt} + \xi_{jt}]}{\exp[x_{lt}\beta - \alpha p_{lt} + \xi_{lt}]}$$

Illustrating Benefits of BLP - Reminder of Problems with Logit

- The own price elasticity e_{jj} was increasing in price (absolute value)
- The cross-elasticity e_{jk}, $k \neq j$ depended only on market share and price of k (but not j!) but not on similarities between goods (IIA)

BLP: No IIA

- There is no IIA at the aggregate (market) level:

$$\frac{s_j(\delta_t, x_t, \theta_2)}{s_l(\delta_t, x_t, \theta_2)} = \frac{\int \frac{\exp[\delta_{jt}(\theta_2) + x_{jt}\tilde{\beta}_i]}{1 + \sum_k \exp[\delta_{kt}(\theta_2) + x_{kt}\tilde{\beta}_i]} dF_{\tilde{\beta}}(\tilde{\beta}_i | \theta_2)}{\int \frac{\exp[\delta_{lt}(\theta_2) + x_{lt}\tilde{\beta}_i]}{1 + \sum_k \exp[\delta_{kt}(\theta_2) + x_{kt}\tilde{\beta}_i]} dF_{\tilde{\beta}}(\tilde{\beta}_i | \theta_2)}$$

- The ratio of market shares depends on the price and characteristics of all the other products

BLP: Price Elasticities

- Finally, using the predicted market shares, the price elasticities are

$$e_{jk} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int \alpha s_{ijt}(1 - s_{ijt}) dF_{\tilde{\beta}}(\tilde{\beta}_i|\theta_2) & \text{if } j = k \\ \frac{p_{kt}}{s_{jt}} \int \alpha s_{ijt} s_{ikt} dF_{\tilde{\beta}}(\tilde{\beta}_i|\theta_2) & \text{otherwise,} \end{cases} \tag{14}$$

where $s_{ijt} = \frac{\exp[\delta_{jt}(\theta_2)+x_{jt}\tilde{\beta}_i]}{1+\sum_k \exp[\delta_{kt}(\theta_2)+x_{kt}\tilde{\beta}_i]}$

BLP: Price Elasticities (Nevo, 2000)

- The price elasticities depend on the density of unobserved consumer types
- Each individual will have a different price sensitivity, which will be averaged to a mean price sensitivity using the individual specific probabilities of purchase as weights
- The price elasticity will be different for different brands
- So if, for example, consumers of BMW have low price sensitivity, then the own-price elasticity of BMW will be low despite the high prices

BLP: Price Elasticities (Nevo, 2000)

- Therefore, substitution patterns are not driven by functional form, but by the differences in the price sensitivity
- The full model also allows for flexible substitution patterns, which are not constrained by a priori segmentation of the market (vs nested logit)

Part 1: 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
- Widely used in IO but also in other fields:
  - Economics is about making choices (demand and supply)
  - Choices differ, usually in some unobservable ways
- Computationally demanding, progress on these challenges
- Many variations of the model in the literature