Lecture 25: Endogeneity III

ResEcon 703: Topics in Advanced Econometrics

Matt Woerman
University of Massachusetts Amherst

Agenda

Last time

- Berry, Levinsohn, and Pakes (1995)

Today

- Berry, Levinsohn, and Pakes (1995)

Upcoming

- Final exam
  - Final exam is posted, due December 19
- Course surveys
  - Course surveys are open, due December 22
  - Click here (`owl.umass.edu/partners/courseEvalSurvey/uma`)

Berry, Levinsohn, and Pakes (1995)

Summary of BLP Theoretical Models

BLP develop models of demand and supply

- Random coefficients model of demand for differentiated products
- Nash-Bertrand pricing model with oligopolistic competition

Both of these models suffer from endogeneity

- BLP need instruments to recover unbiased parameter estimates

The endogeneity enters these models nonlinearly

- BLP develop a new estimation method to handle instruments in a nonlinear model

Instruments

BLP need to find instruments for both the demand equation and the pricing equation

Appropriate instruments must satisfy two conditions

- Correlated with observed product attributes, x_{jt}, and cost characteristics, w_{jt}
- Not correlated with unobserved utility, ξ_{jt}, or unobserved marginal costs, ω_{jt}

BLP construct instruments from the attributes and cost characteristics of all products

- The justification for these instruments comes from a mean independence assumption

Mean Independence Assumption

BLP assume that unobserved utility, ξ_{jt}, and unobserved marginal costs, ω_{jt}, for each product is mean independent of all product attributes and cost characteristics

$$0 = E(\xi_{jt} \mid z_t) = E(\omega_{jt} \mid z_t)$$

where

$$\begin{aligned} z_{jt} &= (x_{jt}, w_{jt}) \\ z_t &= (z_{1t}, \ldots, z_{Jt}) \end{aligned}$$

In other words, observed product attributes and cost characteristics are exogenously (or at least previously) determined, and the unobserved utility and marginal costs are effectively random noise

- But what if all product attributes are jointly determined?

BLP Instruments

BLP argue there are three first-order instruments for each product attribute or cost characteristic

- The attribute (cost characteristic) itself
- The sum of the attribute (cost characteristic) for the firm's other products
- The sum of the attribute (cost characteristic) for the products produced by other firms

They provide some intuition for these instruments, as well as a more formal proof

- Products that face good substitutes tend to have low markups, and products without good substitutes tend to have higher markups
- Markups will respond differently to own and rival products

This argument relies on the exogeneity of attributes in the product space

BLP Estimation

The intuition of the BLP estimation strategy is relatively simple, but practically implementing it is more complex

- Given data on prices and other product attributes, any set of
  1. Observed market shares, S
  2. Demand model parameters, θ_1
  3. Parameters relating to consumer characteristics, θ_2
  4. Cost function parameters, γ

  yields a unique set of unobserved utilities, $\xi_{jt}(x_t, p_t, S_t; \theta_1, \theta_2)$, and unobserved marginal costs, $\omega_{jt}(x_t, p_t, w_{jt}, S_t; \theta_2, \gamma)$
- Assuming we can compute these values, $\xi_{jt}(x_t, p_t, S_t; \theta_1, \theta_2)$ and $\omega_{jt}(x_t, p_t, w_{jt}, S_t; \theta_2, \gamma)$ are uncorrelated with the instruments at the true parameter values
- Use MSM to find the set of parameters that results in no (or minimal) correlation between $\xi_{jt}(x_t, p_t, S_t; \theta_1, \theta_2)$ and $\omega_{jt}(x_t, p_t, w_{jt}, S_t; \theta_2, \gamma)$ and the instruments

But computing $\xi_{jt}(x_t, p_t, S_t; \theta_1, \theta_2)$ and $\omega_{jt}(x_t, p_t, w_{jt}, S_t; \theta_2, \gamma)$ is difficult!

BLP Estimation Steps

BLP use MSM to loop over parameters, θ, and converge to the set of $\hat{\theta}$ that minimize the MSM objective function, and each loop requires four steps that depend on θ

1. Invert the estimation of market shares, $s_{jt}(x_t, p_t, \delta_t; \theta_2)$, to get estimates of $\delta_{jt}(x_t, p_t, S_t; \theta_2)$
2. Solve for the vector of unobserved utilities, $\xi_{jt}(x_t, p_t, S_t; \theta_1, \theta_2)$, implied by $\delta_{jt}(x_t, p_t, S_t; \theta_2)$ and other demand model parameters, θ_1
3. Calculate the vector of unobservable marginal costs, $\omega_{jt}(x_t, p_t, w_{jt}, S_t; \theta_2, \gamma)$, implied by markups, $b_{jt}(x_t, p_t, \delta_t; \theta_2)$, and cost function parameters, γ
4. Calculate all 34 (!) empirical moments, $G(\theta; x, p, w, S)$
    - Product of $\xi_{jt}(x_t, p_t, S_t; \theta_1, \theta_2)$ and 15 demand instruments
    - Product of $\omega_{jt}(x_t, p_t, w_{jt}, S_t; \theta_2, \gamma)$ and 19 cost instruments

BLP Estimation with Logit Assumptions

BLP first estimate their model with the logit assumptions

- $\mu_{ijt} = 0$: no individual-specific utility from observable attributes
- $\varepsilon_{ijt} \sim$ i.i.d. type I extreme value

Estimation with logit assumptions serves two purposes

- Provides a benchmark to compare their random coefficient estimates to demonstrate what is gained by their innovative approach
- Eases the reader into an easier estimation procedure

BLP Logit Model: Step 1

With the logit assumptions, the utility model simplifies to

$$u_{ijt} = \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \varepsilon_{ijt}$$

where

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

Then market shares are simply

$$s_{jt}(x_t, p_t, \delta_t) = \frac{e^{\delta_{jt}}}{1 + \sum_{\ell=1}^{J_t} e^{\delta_{\ell t}}}$$

which is easily inverted to estimate δ_{jt} using observed market shares

$$\delta_{jt}(S_t) = \ln(S_{jt}) - \ln(S_{0t})$$

BLP Logit Model: Steps 2–4

Step 2: The vector of unobserved utilities is easily calculated as

$$\begin{aligned}\xi_{jt}(x_t, p_t, S_t; \theta_1) &= \delta_{jt}(S_t) - x_{jt}\beta + \alpha p_{jt} \\ &= \ln(S_{jt}) - \ln(S_{0t}) - x_{jt}\beta + \alpha p_{jt}\end{aligned}$$

Step 3: The calculation of $\omega_{jt}(x_t, p_t, w_{jt}, S_t; \gamma)$ is also straightforward

Step 4: Empirical moments are calculated as

$$G(\theta; x, p, w, S) = \frac{1}{\sum_{t=1}^{T} J_t} \sum_{t=1}^{T} \sum_{j=1}^{J_t} H_{jt}(z_t) \begin{pmatrix} \xi_{jt}(x_t, p_t, S_t; \theta_1) \\ \omega_{jt}(x_t, p_t, w_{jt}, S_t; \gamma) \end{pmatrix}$$

where $H_{jt}(z_t)$ is the matrix of instruments for product j in market t

BLP uses GMM to find the set of parameters $\hat{\theta}$ that minimizes these empirical moments

BLP Estimation with Mixed Logit Assumptions

The full BLP model uses the less restrictive assumptions of the mixed logit model

- These assumptions yield more realistic substitution patterns between automobiles
- They also make the estimation much more difficult

What makes it so difficult?

- Steps 1 and 3 no longer have closed-form expressions and require simulation
- The inversion in step 1 is computationally infeasible using traditional (at the time) estimation methods

BLP must develop their own estimation method

BLP Mixed Logit Model: Step 1

With the mixed logit assumptions, market shares are

$$s_{jt}(x_t, p_t, \delta_t; \theta_2) = \int L_{ijt}(x_t, p_t, \delta_t, \nu_i, D_i; \theta_2) dP_\nu^*(\nu) d\hat{P}_D^*(D)$$

where

$$L_{ijt}(x_t, p_t, \delta_t, \nu_i, D_i; \theta_2) = \frac{e^{\delta_{jt} + \mu_{ijt}(x_{jt}, p_{jt}, \nu_i, D_i; \theta_2)}}{1 + \sum_{\ell=1}^{J_t} e^{\delta_{\ell t} + \mu_{i\ell t}(x_{\ell t}, p_{\ell t}, \nu_i, D_i; \theta_2)}}$$

These market shares can be simulated as

$$\check{s}_{jt}(x_t, p_t, \delta_t; \theta_2) = \frac{1}{R} \sum_{i=1}^{R} \frac{e^{\delta_{jt} + \mu_{ijt}(x_{jt}, p_{jt}, \nu_i, D_i; \theta_2)}}{1 + \sum_{\ell=1}^{J_t} e^{\delta_{\ell t} + \mu_{i\ell t}(x_{\ell t}, p_{\ell t}, \nu_i, D_i; \theta_2)}}$$

where each ν_i and D_i is drawn from $P_\nu^*(\nu)$ and $\hat{P}_D^*(D)$, respectively

BLP Contraction Mapping

BLP could theoretically estimate the vector of unobserved utilities, $\delta_{jt}(x_t, p_t, S_t; \theta_2)$, by iterating through possible vectors and finding one that yields the observed market shares

- But this is computationally infeasible!

Instead, BLP prove there is a contraction mapping between market shares and $\delta_{jt}(x_t, p_t, S_t; \theta_2)$

- Every possible vector of market shares yields a unique vector of $\delta_{jt}(x_t, p_t, S_t; \theta_2)$
- An iterative algorithm will converge to this unique vector of $\delta_{jt}(x_t, p_t, S_t; \theta_2)$

$$\delta_t^{h+1} = \delta_t^h + \ln(S_t) - \ln(\check{s}_{jt}(x_t, p_t, \delta_t^h; \theta_2))$$

with $\check{s}_{jt}(x_t, p_t, \delta_t^h; \theta_2)$ simulated as on the previous slide

BLP Mixed Logit Model: Steps 2–4

Step 2: The vector of unobserved utilities is easily calculated as

$$\xi_{jt}(x_t, p_t, S_t; \theta_1, \theta_2) = \delta_{jt}(x_t, p_t, S_t; \theta_2) - x_{jt}\beta + \alpha p_{jt}$$

Step 3: The calculation of $\omega_{jt}(x_t, p_t, w_{jt}, S_t; \gamma)$ requires simulation, which is not trivial but also not particularly innovative

- See the paper for more details

Step 4: Empirical moments are calculated as

$$G(\theta; x, p, w, S) = \frac{1}{\sum_{t=1}^{T} J_t} \sum_{t=1}^{T} \sum_{j=1}^{J_t} H_{jt}(z_t) \begin{pmatrix} \xi_{jt}(x_t, p_t, S_t; \theta_1) \\ \omega_{jt}(x_t, p_t, w_{jt}, S_t; \gamma) \end{pmatrix}$$

where $H_{jt}(z_t)$ is the matrix of instruments for product j in market t

BLP uses MSM to find the set of parameters $\hat{\theta}$ that minimizes these empirical moments

Data

BLP use data on cars sold in the US in the years 1971 to 1990

- 20 years of data with 2217 total model-years
- 997 distinct models once similar model-years are aggregated

Variables in each model

- Demand model
  - Constant
  - Horsepower / weight
  - Air conditioning
  - Miles per dollar
  - Size
  - Price
- Supply model
  - Constant
  - Horsepower / weight
  - Air conditioning
  - Miles per gallon
  - Size
  - Trend

Estimation Results

TABLE III

RESULTS WITH LOGIT DEMAND AND MARGINAL COST PRICING (2217 OBSERVATIONS)

Variable	OLS Logit Demand	IV Logit Demand	OLS ln (*price*) on *w*
Constant	−10.068	−9.273	1.882
	(0.253)	(0.493)	(0.119)
*HP / Weight**	−0.121	1.965	0.520
	(0.277)	(0.909)	(0.035)
Air	−0.035	1.289	0.680
	(0.073)	(0.248)	(0.019)
MP$	0.263	0.052	—
	(0.043)	(0.086)	
*MPG**	—	—	−0.471
			(0.049)
*Size**	2.341	2.355	0.125
	(0.125)	(0.247)	(0.063)
Trend	—	—	0.013
			(0.002)
Price	−0.089	−0.216	—
	(0.004)	(0.123)	
No. Inelastic Demands	1494	22	*n.a.*
(+/−2 *s.e.'s*)	(1429–1617)	(7–101)	
R^2	0.387	*n.a.*	.656

Notes: The standard errors are reported in parentheses.
* The continuous product characteristics—hp/weight, size, and fuel efficiency (*MP$* or *MPG*)—enter the demand equations in levels, but enter the column 3 price regression in natural logs.

TABLE IV

ESTIMATED PARAMETERS OF THE DEMAND AND PRICING EQUATIONS: BLP SPECIFICATION, 2217 OBSERVATIONS

Demand Side Parameters	Variable	Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
Means ($\bar{\beta}$'s)	*Constant*	−7.061	0.941	−7.304	0.746
	HP / Weight	2.883	2.019	2.185	0.896
	Air	1.521	0.891	0.579	0.632
	MP$	−0.122	0.320	−0.049	0.164
	Size	3.460	0.610	2.604	0.285
Std. Deviations (σ_β's)	*Constant*	3.612	1.485	2.009	1.017
	HP / Weight	4.628	1.885	1.586	1.186
	Air	1.818	1.695	1.215	1.149
	MP$	1.050	0.272	0.670	0.168
	Size	2.056	0.585	1.510	0.297
Term on Price (α)	$\ln(y-p)$	43.501	6.427	23.710	4.079
Cost Side Parameters					
	Constant	0.952	0.194	0.726	0.285
	ln (*HP / Weight*)	0.477	0.056	0.313	0.071
	Air	0.619	0.038	0.290	0.052
	ln (*MPG*)	−0.415	0.055	0.293	0.091
	ln (*Size*)	−0.046	0.081	1.499	0.139
	Trend	0.019	0.002	0.026	0.004
	$\ln(q)$			−0.387	0.029

Attribute Elasticities

TABLE V

A SAMPLE FROM 1990 OF ESTIMATED DEMAND ELASTICITIES WITH RESPECT TO ATTRIBUTES AND PRICE (BASED ON TABLE IV (CRTS) ESTIMATES)

Model	*HP/Weight*	*Air*	*MP $*	*Size*	*Price*
	Value of Attribute/Price Elasticity of demand with respect to:				
Mazda323	0.366	0.000	3.645	1.075	5.049
	0.458	0.000	1.010	1.338	6.358
Sentra	0.391	0.000	3.645	1.092	5.661
	0.440	0.000	0.905	1.194	6.528
Escort	0.401	0.000	4.022	1.116	5.663
	0.449	0.000	1.132	1.176	6.031
Cavalier	0.385	0.000	3.142	1.179	5.797
	0.423	0.000	0.524	1.360	6.433
Accord	0.457	0.000	3.016	1.255	9.292
	0.282	0.000	0.126	0.873	4.798
Taurus	0.304	0.000	2.262	1.334	9.671
	0.180	0.000	−0.139	1.304	4.220
Century	0.387	1.000	2.890	1.312	10.138
	0.326	0.701	0.077	1.123	6.755
Maxima	0.518	1.000	2.513	1.300	13.695
	0.322	0.396	−0.136	0.932	4.845
Legend	0.510	1.000	2.388	1.292	18.944
	0.167	0.237	−0.070	0.596	4.134
TownCar	0.373	1.000	2.136	1.720	21.412
	0.089	0.211	−0.122	0.883	4.320
Seville	0.517	1.000	2.011	1.374	24.353
	0.092	0.116	−0.053	0.416	3.973
LS400	0.665	1.000	2.262	1.410	27.544
	0.073	0.037	−0.007	0.149	3.085
BMW 735i	0.542	1.000	1.885	1.403	37.490
	0.061	0.011	−0.016	0.174	3.515

Notes: The value of the attribute or, in the case of the last column, price, is the top number and the number below it is the elasticity of demand with respect to the attribute (or, in the last column, price.)

Price Elasticities

TABLE VI

A SAMPLE FROM 1990 OF ESTIMATED OWN- AND CROSS-PRICE SEMI-ELASTICITIES: BASED ON TABLE IV (CRTS) ESTIMATES

	Mazda 323	Nissan Sentra	Ford Escort	Chevy Cavalier	Honda Accord	Ford Taurus	Buick Century	Nissan Maxima	Acura Legend	Lincoln Town Car	Cadillac Seville	Lexus LS400	BMW 735i
323	−125.933	1.518	8.954	9.680	2.185	0.852	0.485	0.056	0.009	0.012	0.002	0.002	0.000
Sentra	0.705	−115.319	8.024	8.435	2.473	0.909	0.516	0.093	0.015	0.019	0.003	0.003	0.000
Escort	0.713	1.375	−106.497	7.570	2.298	0.708	0.445	0.082	0.015	0.015	0.003	0.003	0.000
Cavalier	0.754	1.414	7.406	−110.972	2.291	1.083	0.646	0.087	0.015	0.023	0.004	0.003	0.000
Accord	0.120	0.293	1.590	1.621	−51.637	1.532	0.463	0.310	0.095	0.169	0.034	0.030	0.005
Taurus	0.063	0.144	0.653	1.020	2.041	−43.634	0.335	0.245	0.091	0.291	0.045	0.024	0.006
Century	0.099	0.228	1.146	1.700	1.722	0.937	−66.635	0.773	0.152	0.278	0.039	0.029	0.005
Maxima	0.013	0.046	0.236	0.256	1.293	0.768	0.866	−35.378	0.271	0.579	0.116	0.115	0.020
Legend	0.004	0.014	0.083	0.084	0.736	0.532	0.318	0.506	−21.820	0.775	0.183	0.210	0.043
TownCar	0.002	0.006	0.029	0.046	0.475	0.614	0.210	0.389	0.280	−20.175	0.226	0.168	0.048
Seville	0.001	0.005	0.026	0.035	0.425	0.420	0.131	0.351	0.296	1.011	−16.313	0.263	0.068
LS400	0.001	0.003	0.018	0.019	0.302	0.185	0.079	0.280	0.274	0.606	0.212	−11.199	0.086
735*i*	0.000	0.002	0.009	0.012	0.203	0.176	0.050	0.190	0.223	0.685	0.215	0.336	−9.376

Note: Cell entries i, j, where i indexes row and j column, give the percentage change in market share of i with a \$1000 change in the price of j.

Substitution to the Outside Good

TABLE VII

SUBSTITUTION TO THE OUTSIDE GOOD

	Given a price increase, the percentage who substitute to the outside good (as a percentage of all who substitute away.)	
Model	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

Markups

TABLE VIII

A SAMPLE FROM 1990 OF ESTIMATED PRICE-MARGINAL COST MARKUPS AND VARIABLE PROFITS: BASED ON TABLE 6 (CRTS) ESTIMATES

	Price	Markup Over MC $(p - MC)$	Variable Profits (in \$'000's) $q * (p - MC)$
Mazda 323	\$5,049	\$ 801	\$18,407
Nissan Sentra	\$5,661	\$ 880	\$43,554
Ford Escort	\$5,663	\$1,077	\$311,068
Chevy Cavalier	\$5,797	\$1,302	\$384,263
Honda Accord	\$9,292	\$1,992	\$830,842
Ford Taurus	\$9,671	\$2,577	\$807,212
Buick Century	\$10,138	\$2,420	\$271,446
Nissan Maxima	\$13,695	\$2,881	\$288,291
Acura Legend	\$18,944	\$4,671	\$250,695
Lincoln Town Car	\$21,412	\$5,596	\$832,082
Cadillac Seville	\$24,353	\$7,500	\$249,195
Lexus LS400	\$27,544	\$9,030	\$371,123
BMW 735*i*	\$37,490	\$10,975	\$114,802

Announcements

Office hours

- Reminder: 2:00–3:00 on Tuesdays in 218 Stockbridge

Upcoming

- Final exam is posted, due December 19
- Course surveys are open, due December 22
  - Click here (`owl.umass.edu/partners/courseEvalSurvey/uma`)