Week 3: Empirical Models of Competition with Differentiated Products II

Ben Handel

UC-Berkeley Econ 220A

February 3, 2021

Berry, Levinsohn, Pakes (1995)

Overview

- Seminal paper integrating flexible demand specification with supply side analyzes equilibrium in diff. product auto industry to estimate demand / cost params
- Estimates obtained with only widely available product-level data and aggregate consumer-level data
- Detailed estimates of own and cross price elasticities
- Main primitives in the model are:
  - Joint distribution of consumer characteristics and preferences over attributes/products
  - Price taking consumers
  - Cost function
  - Nash assumption on interactions

Berry, Levinsohn, Pakes (1995)

Overview

- BLP builds on characteristic discrete choice approach and integrates supply conditions to address (at least) two key problems in prior work
- Imposed functional form and restrictive cross-price elasticities:
  - Interacts consumer and product characteristics to yield plausible substitution patterns
- Correlation between prices (observed by econometrician) and product characteristics that may be unobserved to econometrician
  - Similar to 2SLS problem with supply and demand but for differentiated products
  - Makes estimation hard, novel methods
  - Strong mean independence / orthogonality assumption

Functional Forms and Substitution

Berry, Levinsohn, Pakes (1995)

- Random-coefficients, starting point for individual and product characteristic interaction

$$U(\zeta_i, p_j, x_j, \xi_j; \theta) = x_j\overline{\beta} - \alpha p_j + \xi_j + \Sigma_k \sigma_k x_{jk} v_{ik} + \varepsilon_{ij}$$

- $(\zeta_i, \varepsilon_i) = (v_{i1}, v_{i2}, ... v_{iK}, \varepsilon_{i0}, ... \varepsilon_{iK}$ is mean zero vector of random variables with known distribution function
- Now, the contribution of attribute x_k to utility is $(\overline{\beta_k} + \sigma_k v_{ik})x_k$
- Scale variance of v to 1 so that mean and variance of preferences for x_k depend on $\overline{\beta_k}$ and σ_k
- Specification is particularly tractable if ε i.i.d. logit, though integral computation still non-trivial

Functional Forms and Substitution

Berry, Levinsohn, Pakes (1995)

- Now, δ_j is still mean of population utility, but deviation interacts product and consumer characteristics:

$$\mu_{ij} = \Sigma_k \sigma_k x_{jk} v_{ik} + \varepsilon_{ij}$$

- Substitution patterns can now be much more realistic: if price goes up for a large car, consumers are more likely to switch to another large car
- Though random coefficients solves substitution pattern, they prefer specification that makes it easy to incorporate prior info and on functional form for interaction
- Nest random coefficients specification into Cobb Douglas utility specification

Functional Forms and Substitution

Berry, Levinsohn, Pakes (1995)

- Cobb-Douglas random-coefficients utility specification:

$$U(\zeta_i, p_j, x_j, \xi_j; \theta) = (y_i - p_j)^{\alpha} G(x_j, \xi_j, v_i)^{\varepsilon_{ij}}$$

- Empirically, they assume G is linear in logs so that, taking logs:

$$\begin{aligned} u_{ij} &= \alpha log(y_i - p_j) + x_j \overline{\beta} + \xi_j + \Sigma_k \sigma_k x_{jk} v_{ik} + \varepsilon_{ij} \\ u_{i0} &= \alpha log(y_i) + \xi_0 + \sigma_0 v_{i0} + \varepsilon_{i0} \end{aligned}$$

- Income y now interacts with characteristics as well (p) and can be brought in from outside data

- Flexible substitution, income leads to more precise estimates

Endogenous Prices

Berry, Levinsohn, Pakes (1995)

- If producers and consumers know ξ then prices are likely correlated with this unobservable
- Differentiated products analog to classic simultaneity problem
- Complicated by both discrete choice set and interaction of individual and product characteristics
- Models with mean zero independent errors rejected in past studies
- ξ here is analog to endogenous error in homogenous good 2SLS models
  - Can represent, e.g., style, prestige, reputation, past experience
  - ξ alleviates overfitting
- Same as 2SLS, require exogeneity assumption that ξ is mean independent of the instruments

Supply

Berry, Levinsohn, Pakes (1995)

- On supply side, F firms which produce subset of J products each
- Log-linear and CRS assumptions relaxed in robustness

$$ln(mc_j) = w_j\gamma + \omega_j$$

- Observed w_j and unobserved ω_j, could be related to demand attributes
- Profits:

$$\Pi_f = \Sigma_{j\in F_j}(p_j - mc_j)Ms_j(p, x, \xi; \theta)$$

- J FOC for Nash-Bertrand pricing model imply markups:

$$s_j(p, x, \xi; \theta) + \Sigma_{r\in F_j}(p_r - mc_r)\frac{\partial s_r(p, x, \xi; \theta)}{\partial p_j} = 0$$

Supply

Berry, Levinsohn, Pakes (1995)

- Define matrix Δ_{jr} to describe cross-price derivatives for multi-product firms, similar to H in Bresnahan
- Putting into matrix notation, FOC is now:

$$p = mc + \Delta(p, x, \xi; \theta)^{-1} s(p, x, \xi; \theta)$$

- Second part of this term is the markup. Depends on demand system parameters and equilibrium price
- Since p is function of ω need to substitute in to take to data:

$$ln(p - b(p, x, \xi; \theta)) = w\gamma + \omega$$

- Need appropriate instruments for ω

Instruments

Berry, Levinsohn, Pakes (1995)

- Appropriate instruments are correlated with observed demand and cost characteristics but not correlated with demand (or supply) disturbances ξ and ω
- Mean independence assumption is central assumption here:

$$\begin{aligned} E[\xi_j|z] &= E[\omega_j|z] = 0 \\ z_j &= [x_j w_j] \end{aligned}$$

- Intuition for observed characteristics of all products: if face similar products it implies different markup patterns without correlation with ξ, cost shifters serve usual role in shifting supply.
- What happens if there is a true underlying model that determines observed and unobserved characteristics?

Estimation Overview

Berry, Levinsohn, Pakes (1995)

- New methods for (i) random coefficients and (ii) endogeneity in characteristics-based discrete choice model
- In theory, given data on p_j and x_j, any choice of triple with (1) observed market shares (2) consumers characteristic distribution and (3) model parameters implies unique sequences ξ_j and ω_j
- Assuming can compute these values (half of the paper!):
  - Mean independent moment restrictions imply z uncorrelated with results evaluated at $\theta = \theta_0$
  - GMM: find θ that sets instruement equations as close as possible to 0
  - Intuition: errors should be mean 0 conditional on instruments, think of OLS
- 34 instruments, 15 demand, 19 cost, and 34 moment equations

Computation

Berry, Levinsohn, Pakes (1995)

- Illustrations for logit and mixed logit models
- Key steps to evaluate $G_J(\theta, s_n, P_{ns})$ moment conditions:
  - Estimate market shares implied by model
  - Solve for ξ implied by simulated and observed market shares
  - Calculate ω from difference in price / markups computed given shares
  - Calculate instruments and compute objective function
- See computation in BLP section 6 and Nevo: look at logit first
- Random coefficients process more complex, requires integral simulation

Data

Berry, Levinsohn, Pakes (1995)

- Product Characteristics from annual issue of trade book for assumed base models from 1971-1990:
  - Number of cylinders
  - Number of doors
  - Weight
  - Horsepower
  - Length
  - Width
  - MPG
  - Air conditioning
- Treating model/year as observation they have 2217 observations
- Bring in price of gasoline, market size, and reliability ratings
- Firm multi-product description

Data Description

Berry, Levinsohn, Pakes (1995)

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.

Results

Berry, Levinsohn, Pakes (1995)

- Computational concerns dictate few attributes:
  - Horsepower to weight ratio
  - Air conditioning dummy
  - Miles per dollar
  - Size
- They start by comparing:
  - OLS Logit
  - IV Logit

Results

Berry, Levinsohn, Pakes (1995)

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.

- Implies, not surprisngly, strange markup patterns

Full Model Results

Berry, Levinsohn, Pakes (1995)

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

Full Model Results

Berry, Levinsohn, Pakes (1995)

TABLE V

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

Model	Value of Attribute/Price Elasticity of demand with respect to: *HP/Weight*	*Air*	*MP $*	*Size*	*Price*
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.)

Full Model Results

Berry, Levinsohn, Pakes (1995)

- Elasticity with respect to MP$ illustrates the importance of thinking about both the mean and variance of the consumer taste distribution
- Elasticity wrt MP$ declines monotonically with MP$ rating
- 10% increase in MP$ increases demand for Mazda 323, Sentra and other small cars it has no effect for cars with low *MP$*
- People in high end of distribution sbustituting around small cars quickly, but no one in big cars really cares

Key Contribution: Elasticities

Berry, Levinsohn, Pakes (1995)

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.

Outside Good Substitution

Berry, Levinsohn, Pakes (1995)

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

Estimated Markups

Berry, Levinsohn, Pakes (1995)

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

Estimated Markups

Berry, Levinsohn, Pakes (1995)

TABLE IX

RESULTS FROM SOME ALTERNATIVE SPECIFICATIONS:
PRICE-MARGINAL COST MARKUPS

	Base Case (reported in Table VIII)	Include ln (q) in cost function	Use *AT* instead of *AIR*	*Weight and HP* instead of *HP/Wt*	Include interaction terms in cost function	Use 3 region dummies and add *Reliability*	Add *weight* and include interactions in cost function
Mazda 323	$ 801	$1,616	$1,012	$1,073	$ 828	$1,125	$1,389
Nissan Sentra	$ 880	$1,769	$1,153	$1,271	$ 912	$1,308	$1,487
Ford Escort	$1,077	$2,043	$1,326	$1,470	$1,111	$2,094	$1,690
Chevy Cavalier	$1,302	$2,490	$1,729	$1,655	$1,329	$2,593	$2,020
Honda Accord	$1,992	$3,059	$2,629	$2,703	$2,059	$3,839	$2,327
Ford Taurus	$2,577	$3,721	$2,528	$3,344	$2,585	$4,094	$2,898
Buick Century	$2,420	$4,162	$3,161	$2,939	$2,405	$4,030	$3,321
Nissan Maxima	$2,881	$4,674	$4,565	$2,085	$2,911	$6,941	$3,513
Acura Legend	$4,671	$7,105	$6,563	$3,059	$4,661	$8,305	$5,081
Lincoln Town Car	$5,596	$8,029	$6,778	$4,765	$5,508	$7,114	$6,518
Cadillac Seville	$7,500	$10,733	$8,635	$4,863	$7,439	$9,182	$8,015
Lexus LS400	$9,030	$10,510	$8,411	$4,791	$8,585	$10,925	$7,398
BMW 735i	$10,975	$13,646	$9,122	$7,605	$10,713	$12,153	$12,202
No. of demand side variables significant at 95% level[a]	5 of 5	4 of 5	4 of 5	6 of 6	4 of 5	8 of 8	4 of 6

[a] A demand side variable is considered significant if *either* its mean or standard deviation (σ_β) is significant. See text for details.

Empirical Purpose: Realistic Equilibrium?

Berry, Levinsohn, Pakes (1995)

- Paper is not testing a specific behavioral assumption or making a specific point about conduct in the auto industry
- Instead, it seeks to generally show the equilibrium results seems ‘more reasonable’ with novel methods:
  - They show that endogeneity matters
  - They show that flexible substitution patterns matter
- They could integrate their model with different pricing / supply assumptions a la Bresnahan (1987)

Applications

Berry, Levinsohn, Pakes (1995)

- Broad types of applications:
  - Conditional on primitives, model can solve for distribution of prices, quantities, variable profits, consumer welfare
  - Investigate changes to primitives assuming others fixed
  - Change that primitives played in historical movements in data

- Potential specific applications:
  - Trade policy
  - Merger policy
  - Environmental policy

- Counterfactuals and welfare *subject to otherwise similar environments*

Extensions

Berry, Levinsohn, Pakes (1995)

- Realistic treatment of endogenous product / dynamics from firm perspective
- Realistic treatment of dynamics from consumer perspective. Durable goods and intertemporal substitution matter
  - We will look at Hendel & Nevo (2006) in two weeks
  - What about used cars?
- Incorporating micro-level data to obtain more precise estimates seems important
- No real discussion of identification, quasi-experimental variation would be great: this is a fantastic framework to deal with endogeneity, but it might be better to find better data also!

Discussion

Berry, Levinsohn, Pakes (1995)

- What variation in the data is identifying the random coefficient distribution parameters here?
    - What is the ideal data setting here?
    - What would be different with individual level data about identification? Panel data?
- Large focus on endogeneity and instrumental variable approach. What are the instruments and why do they work? How strong are these instruments?
- What are the potential positives and negatives of using similar methods in your own empirical work if you can't acquire better data?
- What about BLP's complexity is valuable? What are the shortcomings?
    - Without compelling reduced form evidence for a question, will anyone believe your results?