Estimation of Demand Model 3

BLP Approach 2 –Instruments and Applications–

Yuta Toyama

Last updated: December 2, 2018

Introduction

- Last week: Introduce BLP framework, with an emphasis on model and estimation approach.
- This week:
  - Instruments commonly used in BLP framework
  - Applications: BLP (1995, EMA), Nevo (2001, EMA)

Role of IVs

- Estimation of BLP-type model is based on

$$E[\xi_{jt}|z_{jt}] = 0,$$

  where z_{jt} is a vector of instruments.
- Remember that we should satisfy

$$(\#\ of\ \ moments) \geq (\#\ of\ \ parameters)$$

- IV has two roles
  - Deal with price endogeneity: $Cov(p_{jt}, \xi_{jt}) \neq 0$
  - Moment conditions to identify nonlinear parameters θ_2

- Consider a nested logit model

$$\log \frac{s_{jt}}{s_{0t}} = \beta_0 + \alpha p_{jt} + \beta x_{jt} + \sigma \ln s_{jt|g} + \xi_{jt}$$

where x_{jt} is a scalar (for simplicity).

- Endogenous variables
  - p_{jt} due to price endogeneity
  - within-group market share $s_{jt|g} (\equiv s_{jt}/(\sum_{j \in G_g} s_{jt}))$ is endogenous.
  - Note: Typically we assume that product characteristics x_{jt} is exogenous.

- Moment conditions

$$\mathbb{E}\left[\xi_{jt}\left(1, x_{jt}, z_{jt}^1, z_{jt}^2\right)\right] = 0$$

where z_{jt}^1 and z_{jt}^2 are additional instruments.

- You can run GMM based on it.
  - Or, equivalently, you can run the 2SLS by using (z_{jt}^1, z_{jt}^2) as additional instruments.

Overview of IVs

- Instruments for price endogeneity
  1. "So called" BLP instruments
  2. Cost-based instruments (so-called Hausman-Nevo instruments)

- Instruments to identify non-linear parameters
  1. I focus on the case of nested logit.
  2. For the case of random coefficient logit, see Reynaert and Verboven (2014, JOE), Gandhi and Houde (2016, WP)

IV for price 1: Competition in Characteristics Space

- Assumption:

$$E[\xi_{jt}|\{x_{jt}\}_{j\in J_t}] = 0$$

- BLP (1995) proposed
  1. Own characteristics x_{jt}
  2. Sum of char of other products produced by the firm $\sum_{k\in J_f, k\neq j} x_{kt}$ where J_f is the set of own products
  3. Sum of char of competitors' products: $\sum_{k\notin J_f} x_{kt}$

- Relevance of IV: proximity in characteristics space to other products $\rightarrow$ markup $\rightarrow$ price

- Independence: x_{jt} are assumed to be set before ξ_{jt} is known.

- Most commonly used b.c. it does not require additional data.

IV for price 2: Cost-based

- Cost data are rarely directly observed, especially at the product level.
  - Note: Price p_{jt} has variation at product-and-market level.
- Villas-Boas (2007) uses prices of inputs interacted with product dummy variables.
- Hausman (1996) and Nevo (2001)
  - Use prices of the product in **other geographical markets.**
  - Independence: after controlling for common effects, the unobserved characteristics are assumed independent across markets.
  - Relevance: prices will be correlated across markets due to common marginal cost shocks.
  - Easy to come up with examples where IVs are not valid (e.g., national promotions).

IV for nonlinear parameters

- Consider a nested logit model without price endogeneity.

$$\log \frac{s_{jt}}{s_{0t}} = \beta_0 + \beta x_{jt} + \sigma \ln s_{jt|g} + \xi_{jt}$$

- IV for $s_{jt|g}$:
  - sum of characteristics of own products in group g : $\sum_{k \in J_f, k \in G_g, k \neq j} x_{kt}$
  - sum of characteristics of competitors' products in group g : $\sum_{k \notin J_f, k \in G_g, k \neq j} x_{kt}$
- Idea: The within-group market share depends on the degree of competition within group g.

BLP (1995) "Automobile Prices in Market Equilibrium"

- Research Question:
  - Develops a technique for empirically analyzing demand and supply in differentiated products markets
  - Apply it to the U.S. automobile industry
- Take-away:
  - The role of instruments
  - Comparison between (multinomial) logit and RC logit

Data

- 20 years of annual US national data.
  - Period: $1971 - 90$.
  - Sample size: 2217 model-years
- Variables:
  - Quantity data by name plate.
  - List price
  - characteristics from Automotive News Market Data Book

Demand Model

- The indirect utility from product j

$$u_{ijt} = x_{jt}\beta_i + \alpha \log(y_i - p_{jt}) + \xi_{jt} + \epsilon_{ijt}$$

  - y_i is income. This allows income effect in utility function.

- Random coefficient:

$$\beta_i^k = \beta^k + \sigma^k \nu_{ik}, \nu_{ik} \sim N(0,1)$$

- The utility from the outside option ($j = 0$).

$$u_{i0t} = \alpha \log(y_i) + \epsilon_{i0t}$$

  - Note: $\xi_{0t} = 0$ as normalization.

Estimation

- Combine both demand and supply moments
  - supply moments help to pin down demand parameters.

- Demand: $E[z_{jt}^d \xi_{jt}] = 0$, where z_{jt}^d includes
  - own characteristics x_{jt}
  - sum of char of other products by the same firm
  - sum of char of products by other firms.

- Supply side: Differentiated product Bertrand competition.
  - Marginal cost function: $\log(mc_{jt}) = w_{jt}\gamma + \omega_{jt}$
    - w_{jt}: product characteristics (overlapped with x_{jt})
  - mc_{jt} is implied from the FOC with demand model.
  - Use moment condition $E[z_{jt}^s \omega_{jt}] = 0$ where z_{jt}^s include w_{jt} and x_{jt}.

- System GMM by stacking $E[z_{jt}^d \xi_{jt}] = 0$ and $E[z_{jt}^s \omega_{jt}] = 0$

Table 3: Role of IV (in multinomial logit)

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 (+/− 2 *s.e.'s*)	1494 (1429–1617)	22 (7–101)	*n.a.*
R^2	0.387	*n.a.*	.656

Table 4: Full Result

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

Table 5: Elasticity

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

Table 7: Substitution to the outside option

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

Table 8: Markups

	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

Summary

- Powerful method & seminal work that inspired so many applications in the field.
- The paper demonstrates
  - importance of instruments
  - Comparison between RCDC and simple logit.
- Common complaints:
  - Instruments
  - No demand side dynamics.
    - Gowrisankaran and Rysman (2012, JPE) extends the BLP-type demand model into durable goods.

"Measuring Market Power in the Ready-to-eat Cereal Industry"

- Research questions:
  - Is pricing in the industry collusive?
  - What portion of the markup in the indusry due to
    - product differentiation?
    - multi-product firms?
    - Potential price collusion?
- Take-away:
  - effects of various IVs
  - testing the model of competition
  - industry where characteristics are less obvious.

The Ready-to-eat (RTE) cereal industry

- Characterized by:
  - High concentration (C3 75%, C6 90%)
  - High price-cost margins: 45%
  - Large advertising to sales ratios: 13%
  - Numerous introductions of brands: 67 new brands by top 6 in 80's.
- This has been used to claim that this is a perfect example of collusive pricing.

Strategy

- Estimate brand-level demand
- Compute price-cost-margin (PCM) by different industry structures/ models of conduct
  - single-product firms
  - current ownership (multi-product firms)
  - Fully-collusive pricing (joint ownership)
- Compare predicted PCM to observed PCM.

(Recap) Supply

- Firm f produces J_f products. The profit for firm f

$$\pi_f = \sum_{j \in J_f} p_j q_j(\mathbf{p}) - C_j(q_j(\mathbf{p}))$$

- Solution concept: **Bertrand-Nash equilibrium**
- Given prices of products offered by competitors, FOCs are

$$q_j(\mathbf{p}) + \sum_{r \in J_f} (p_r - mc_r) \frac{\partial q_r(\mathbf{p})}{\partial p_j} = 0, \forall j \in J_f$$

or

$$\mathbf{q}(\mathbf{p}) - \underbrace{\Omega(\mathbf{p})}_{(J \times J)} \underbrace{(\mathbf{p} - \mathbf{mc})}_{(J \times 1)} = \mathbf{0}$$

- $\Omega(\mathbf{p}) = \Omega^* * S(\mathbf{p})$,where $*$ is element-by-element product
- Ω^*: ownership matrix. $S(\mathbf{p})$: derivative of market share w.r.t. price.

Demand

- Utility same as before

$$u_{ijt} = x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

  - x_j includes brand dummy variables D_j, which captures the characteristics that does not vary by market.

Data

- IRI Infoscan scanner data
  - Market shares: defined by converting volume to servings
    - Market size: one serving per consumer per day.
  - prices: pre-coupon real transaction per serving price
  - 25 brands
  - 67 cities
  - over 20 quarters (1988-1992)
  - 1124 markets, 27862 observations.
- Demographics of each market from March CPS
- Cost instruments (wage) from Monthly CPS

Estimation

- Use only demand side moments.
- Various instruments:
  - Characteristics of competition (BLP IV)
    - problematic for this sample, because product characteristics are same across markets!
  - Prices in other cities
  - Proxies for city-level costs: density, earning in retail sector, transportation costs.

Table 5: Logit demand with various IVs

	OLS			IV						
Variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Price	−4.96	−7.26	−7.97	−8.17	−17.57	−17.12	−22.56	−23.77	−23.37	−23.07
	(0.10)	(0.16)	(0.15)	(0.11)	(0.50)	(0.49)	(0.51)	(0.53)	(0.47)	(1.17)
Advertising	0.158	0.026	0.026	0.157	0.020	0.020	0.018	0.017	0.018	0.013
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log of Median Income	—	—	0.89	—	—	—	1.06	1.13	1.12	—
			(0.02)				(0.02)	(0.02)	(0.02)	
Log of Median Age	—	—	−0.423	—	—	—	−0.063	0.003	−0.007	—
			(0.052)				(0.059)	(0.062)	(0.061)	
Median HH Size	—	—	−0.126	—	—	—	−0.053	−0.036	−0.038	—
			(0.027)				(0.029)	(0.031)	(0.031)	
Fit/Test of Over Identification[b]	0.54	0.72	0.74	436.9	168.5	181.2	83.96	82.95	85.87	15.06
				(26.30)	(30.14)	(16.92)	(30.14)	(16.92)	(42.56)	(42.56)
1st Stage R^2	—	—	—	0.889	0.908	0.908	0.910	0.909	0.913	0.952
1st Stage F-test	—	—	—	5119	124	288	129	291	144	180
Instruments[c]	—	—	—	brand dummies	prices	cost	prices	cost	prices, cost	prices, cost

Table 6: Distribution of Price Coefficient

Variable	Means (β's)	Standard Deviations (σ's)	Interactions with Demographic Variables: Income	Income Sq	Age	Child
Price	−27.198	2.453	315.894	−18.200	—	7.634
	(5.248)	(2.978)	(110.385)	(5.914)		(2.238)
Advertising	0.020	—	—	—	—	—
	(0.005)					
Constant	−3.592[b]	0.330	5.482	—	0.204	—
	(0.138)	(0.609)	(1.504)		(0.341)	
Cal from Fat	1.146[b]	1.624	—	—	—	—
	(0.128)	(2.809)				
Sugar	5.742[b]	1.661	−24.931	—	5.105	—
	(0.581)	(5.866)	(9.167)		(3.418)	
Mushy	−0.565[b]	0.244	1.265	—	0.809	—
	(0.052)	(0.623)	(0.737)		(0.385)	
Fiber	1.627[b]	0.195	—	—	—	−0.110
	(0.263)	(3.541)				(0.0513)
All-family	0.781[b]	0.1330	—	—	—	
	(0.075)	(1.365)				
Kids	1.021[b]	2.031	—	—	—	
	(0.168)	(0.448)				
Adults	1.972[b]	0.247	—	—	—	
	(0.186)	(1.636)				
GMM Objective (degrees of freedom)			5.05 (8)			
MD χ^2			3472.3			
% of Price Coefficients > 0			0.7			

Figure 2: Logit demand with various IVs

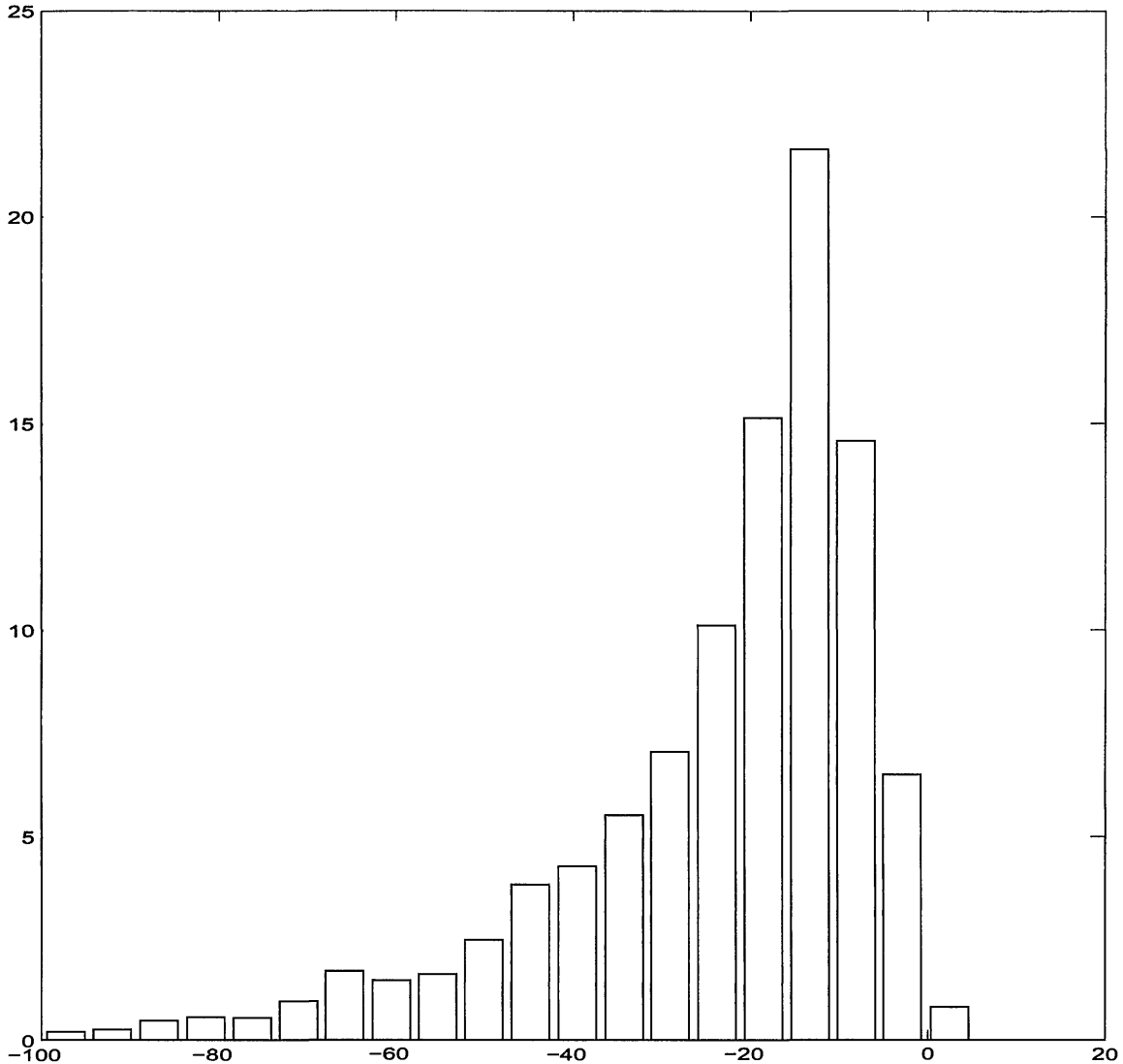

FIGURE 2.—Frequency distribution of price coefficient (based on Table VI).

Table 7: Price Elasticity Matrix

#	Brand	Corn Flakes	Frosted Flakes	Rice Krispies	Froot Loops	Cheerios	Total	Lucky Charms	P Raisin Bran	CapN Crunch	Shredded Wheat
1	K Corn Flakes	−3.379	0.212	0.197	0.014	0.202	0.097	0.012	0.013	0.038	0.028
2	K Raisin Bran	0.036	0.046	0.079	0.043	0.145	0.043	0.037	0.057	0.050	0.040
3	K Frosted Flakes	0.151	−3.137	0.105	0.069	0.129	0.079	0.061	0.013	0.138	0.023
4	K Rice Krispies	0.195	0.144	−3.231	0.031	0.241	0.087	0.026	0.031	0.055	0.046
5	K Frosted Mini Wheats	0.014	0.024	0.052	0.043	0.105	0.028	0.038	0.054	0.045	0.033
6	K Froot Loops	0.019	0.131	0.042	−2.340	0.072	0.025	0.107	0.027	0.149	0.020
7	K Special K	0.114	0.124	0.105	0.021	0.153	0.151	0.019	0.021	0.035	0.035
8	K Crispix	0.077	0.086	0.114	0.034	0.181	0.085	0.030	0.037	0.048	0.043
9	K Corn Pops	0.013	0.109	0.034	0.113	0.058	0.025	0.098	0.024	0.127	0.016
10	GM Cheerios	0.127	0.111	0.152	0.034	−3.663	0.085	0.030	0.037	0.056	0.050
11	GM Honey Nut Cheerios	0.033	0.192	0.058	0.123	0.094	0.034	0.107	0.026	0.162	0.024
12	GM Wheaties	0.242	0.169	0.175	0.025	0.240	0.113	0.021	0.026	0.050	0.043
13	GM Total	0.096	0.108	0.087	0.018	0.131	−2.889	0.017	0.017	0.029	0.029
14	GM Lucky Charms	0.019	0.131	0.041	0.124	0.073	0.026	−2.536	0.027	0.147	0.020
15	GM Trix	0.012	0.103	0.031	0.109	0.056	0.026	0.096	0.024	0.123	0.016
16	GM Raisin Nut	0.013	0.025	0.042	0.035	0.089	0.040	0.031	0.046	0.036	0.027
17	GM Cinnamon Toast Crunch	0.026	0.164	0.049	0.119	0.089	0.035	0.102	0.026	0.151	0.022
18	GM Kix	0.050	0.279	0.070	0.101	0.106	0.056	0.088	0.030	0.149	0.025
19	P Raisin Bran	0.027	0.037	0.068	0.044	0.127	0.035	0.038	−2.496	0.049	0.036
20	P Grape Nuts	0.037	0.049	0.088	0.042	0.165	0.050	0.037	0.051	0.052	0.047
21	P Honey Bunches of Oats	0.100	0.098	0.104	0.022	0.172	0.109	0.020	0.024	0.038	0.033
22	Q 100% Natural	0.013	0.021	0.046	0.042	0.103	0.029	0.036	0.052	0.046	0.029
23	Q Life	0.077	0.328	0.091	0.114	0.137	0.046	0.096	0.023	0.182	0.029
24	Q CapN Crunch	0.043	0.218	0.064	0.124	0.101	0.034	0.106	0.026	−2.277	0.024
25	N Shredded Wheat	0.076	0.082	0.124	0.037	0.210	0.076	0.034	0.044	0.054	−4.252
26	Outside good	0.141	0.078	0.084	0.022	0.104	0.041	0.018	0.021	0.033	0.021

Table 8: Estimates of Markups

- Retail margin in data: 46%.

	Logit (Table V column ix)	Full Model (Table VI)
Single Product Firms	33.6% (31.8%–35.6%)	35.8% (24.4%–46.4%)
Current Ownership of 25 Brands	35.8% (33.9%–38.0%)	42.2% (29.1%–55.8%)
Joint Ownership of 25 Brands	41.9% (39.7%–44.4%)	72.6% (62.2%–97.2%)
Current Ownership of All Brands	37.2% (35.2%–39.4%)	—
Monopoly/Perfect Price Collusion	54.0% (51.1%–57.3%)	—

[a] Margins are defined as $(p - mc)/p$. Presented are medians of the distribution of 27,862 (brand-city-quarter) observations. 95% confidence intervals for these medians are reported in parentheses based on the asymptotic distribution of the estimated demand coefficients. For the Logit model the computation is analytical, while for the full model the computation is based on 1,500 draws from this distribution.