Week 4: Empirical Models of Competition with Differentiated Products II

Ben Handel

UC-Berkeley Econ 220A

February 9, 2022

Logistics

Week 4

- Referee Report 1 due soon
- Problem Set 1 in two weeks
- Research proposals: start thinking!

Ready to Eat Cereal

Nevo (2001)

- **Methodology:** Extends BLP in a few ways that improve empirical applicability. Primary contributions:
  - Brand-specific demand intercepts improve model fit substantially, implies use of completely different instruments
  - Implies need to add extra equations to identify β
  - Model heterogeneity as function of empirical non-parametric demographic distributions

- **Application:** Better data, clearer explication of methods than BLP, specific question similar to Bresnahan (1987):
  - Are cereal manufacturers colluding (market power) or are high margins the result of purely differentiation?
  - Three sources of margins: Product differentiation, multi-product pricing, collusion
  - Collusion in prices or in product offerings?

Context

Nevo (2001)

- Read to Eat Cereal Industry: Three billion pounds of cereal, $9 billion in sales
- High levels of concentration and high profit margins have led to substantial antitrust interest: FTC brought complaint against industry on number of brands
- 18% of expenses spent on advertising
- Actually complicated to produce: several methods
- Data:
  - Market shares and prices in each market (city-quarter)
  - Brand characteristics (size?)
  - Advertising
  - Information on distribution of demographics
- 65 cities and 25 brands per city, 1988-1992

Data Description

Nevo (2001)

TABLE I

VOLUME MARKET SHARES

	88Q1	88Q4	89Q4	90Q4	91Q4	92Q4
Kellogg	41.39	39.91	38.49	37.86	37.48	33.70
General Mills	22.04	22.30	23.60	23.82	25.33	26.83
Post	11.80	10.30	9.45	10.96	11.37	11.31
Quaker Oats	9.93	9.00	8.29	7.66	7.00	7.40
Ralston	4.86	6.37	7.65	6.60	5.45	5.18
Nabisco	5.32	6.01	4.46	3.75	2.95	3.11
C3	75.23	72.51	71.54	72.64	74.18	71.84
C6	95.34	93.89	91.94	90.65	89.58	87.53
Private Label	3.33	3.75	4.63	6.29	7.13	7.60

Source: IRI Infoscan Data Base, University of Connecticut, Food Marketing Center.

Data Description

Nevo (2001)

TABLE II

AGGREGATE ESTIMATES OF PRODUCTION COSTS

	RTE Cereal (SIC 2043)		All Food Industries (SIC 20)	
Item	M$	% of value	M$	% of value
Value of Shipments	8,211	100.0	371,246	100.0
Materials	2,179	26.5	235,306	63.4
Labor	677	8.2	32,840	8.8
Energy	76	0.9	4,882	1.3
Gross Margin		64.4		26.5

Source: Annual Survey of Manufacturers 1988–1992.

Data Description

Nevo (2001)

TABLE III

DETAILED ESTIMATES OF PRODUCTION COSTS

Item	$/lb	% of Mfr Price	% of Retail Price
Manufacturer Price	2.40	100.0	80.0
Manufacturing Cost:	1.02	42.5	34.0
Grain	0.16	6.7	5.3
Other Ingredients	0.20	8.3	6.7
Packaging	0.28	11.7	9.3
Labor	0.15	6.3	5.0
Manufacturing Costs (net of capital costs)[a]	0.23	9.6	7.6
Gross Margin		57.5	46.0
Marketing Expenses:	0.90	37.5	30.0
Advertising	0.31	13.0	10.3
Consumer Promo (mfr coupons)	0.35	14.5	11.7
Trade Promo (retail in-store)	0.24	10.0	8.0
Operating Profits	0.48	20.0	16.0

[a] Capital costs were computed from ASM data.
Source: Cotterill (1996) reporting from estimates in CS First Boston Reports "Kellogg Company," New York, October 25, 1994.

Data Description

Nevo (2001)

TABLE IV

PRICES AND MARKET SHARES OF BRANDS IN SAMPLE

Description	Mean	Median	Std	Min	Max	Brand Variation	City Variation	Quarter Variation
Prices (¢ per serving)	19.4	18.9	4.8	7.6	40.9	88.4%	5.3%	1.6%
Advertising (M$ per quarter)	3.56	3.04	2.03	0	9.95	66.2%	—	1.8%
Share within Cereal Market (%)	2.2	1.6	1.6	0.1	11.6	82.3%	0.5%	0%

Source: IRI Infoscan Data Base, University of Connecticut, Food Marketing Center.

Empirical Framework

Nevo (2001)

- Demand parameters used to compute PCM implied by different models of conduct: compare to observed PCM measures

- Supply equations: same as BLP / Bresnahan

$$
\begin{aligned}
\Pi_f &= \Sigma_{j \in J_f}(p_j - mc_j) M s_j(p) - C_f \\
0 &= s_j(p) + \Sigma_{r \in J_f}(p_r - mc_r) \frac{\partial s_r(p)}{\partial p_j} \\
\Omega_{jr} &= \Omega^*_{jr} S_{jr} \\
p - mc &= \Omega^{-1} s(p)
\end{aligned}
$$

- Estimate PCM without observing actual cost: comes from demand estimates and supply models

Empirical Demand Framework

Nevo (2001)

- Conditional i,j,t indirect utility is:

$$u_{ijt} = x_j \beta_i^* - \alpha_i^* p_{jt} + \xi_j + \Delta\xi_{jt} + \epsilon_{ijt}$$

- Vertical / quality component and market-specific components

- Taste parameter heterogeneity:

$$\begin{bmatrix} \alpha_i^* \\ \beta_i^* \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \Pi D + \Sigma v_i, v_i \to N(0, I_{K+1})$$

- Π measures how tastes vary with demographics, Σ is covariance matrix

Empirical Demand Framework

Nevo (2001)

- Outside good utility normalized to 0:

$$u_{i0t} = \xi_0 + \pi_0 D_i + \sigma_0 v_{i0} + \epsilon_{i0t}$$

- Specification with linear parameters $\theta_1 = (\alpha, \beta)$ and non-linear parameters $\theta_2 = (\Pi, \Sigma)$:

$$u_{ijt} = \delta_j(x_j, p_{jt}, \xi_j, \Delta_{\xi jt}; \theta_1) + \mu_{ijt}(x_j, p_{jt}, v_i, D_i; \theta_2) + \epsilon_{ijt}$$
$$\delta_{jt} = x_j\beta - \alpha p_{jt} + \xi_j + \Delta\xi_{jt},\ \mu_{ijt} = [p_{jt}, x_j]' * (\Pi D_i + \Sigma v_i)$$

- Mean utility and mean-zero heteroskedastic deviation captures random coefficients

Empirical Demand Framework

Nevo (2001)

- Consumers maximize utility in each period
- Implies market share equation:

$$s_{jt}(x, p_{.t}, \delta_{.t}; \theta_2) = \int_{A_{jt}} dP^*(D, v, \epsilon) = \int_{A_{jt}} dP^*(\epsilon) dP^*(v) dP^*(D)$$

- Latter is true because D, v, and ϵ are assumed to be independent

Demand Framework Discussion

Nevo (2001)

- Straightforward way to estimate model would be to choose parameters that minimize distance between predicted market shares and actual market shares
- Actual estimation has to be more complex because of endogeneity
- Logit assumption often made to simplify market share integral
- Nested logit and GEV models allow more flexibility, with a priori segmentation
- Full model allows flexible own and cross price elasticities

Estimation

Nevo (2001)

- Estimation here is similar in spirit to BLP though there are some differences in the details
  - Different instrumental variables / identification assumptions
  - BLP relies on supply equation functional form
  - Brand fixed effects
- Generalized Method of Moments estimation
  - Exploit population moment conditions with instrumental variables and structural error
- Let Z be a vector of instruments such that $E[Z' \cdot \omega(\theta^*)] = 0$
- ω is error term defned shortly, θ^* is true parameter value

Estimation

Nevo (2001)

- The GMM estimate is:

$$\hat{\theta} = arg \min_{\theta} \omega(\theta)' Z A^{-1} Z' \omega(\theta)$$

- Weight matrix A makes estimate more efficient
- Intuition: think about least squares with a more complicated objective function
- Error term $\omega(\theta)$ is unobserved product characteristics $(\xi_j + \Delta\xi_{jt})$ in BLP: here with brand dummies it is just $\Delta\xi_{jt}$
- These are computed to form moments by:
  - Solving for mean utilities that solve implicit market share equations
  - Must invert system of market share equations to find $\delta_{.t}$

Estimation

Nevo (2001)

- System of implicit equations for market share to recover δ from:

$$s_{.t}(x, p_{.t}, \delta_{.t}; \theta_2) = S_{.t}$$

- For the logit model inverting this equation is simple: $\delta_{jt} = ln(S_{jt} - ln(S_{0t})$
- For full model without convenient logit form, this is complicated and has to be done numerically with a contraction mapping
- Once this has been done, error term is defined as:

$$\omega_{jt} = \delta_{jt}(x, p_{.t}, S_{.t}; \theta_2) - (x_j\beta + \alpha p_{jt})$$

- With logit you can just run OLS or IV with 2SLS. With mixed logit, nonlinear terms in δ make this impossible so you do GMM with IV

Estimation

Nevo (2001)

- Numerical inversion to get market shares in random-coefficients logit:

$$s_{jt}(p_{\cdot t}, x_{\cdot t}, \delta_{\cdot t}, P_{ns}; \theta_2)$$
$$= \frac{1}{ns}\sum_{i=1}^{ns} s_{jti} = \frac{1}{\text{ns}}\sum_{i=1}^{\text{ns}}$$
$$\times \frac{\exp\left[\delta_{jt} + \Sigma_{k=1}^{K} x_{jt}^{k}(\sigma_k v_i^k + \pi_{k1}D_{i1} + \cdots + \pi_{kd}D_{id})\right]}{1 + \Sigma_{m=1}^{J} \exp\left[\delta_{mt} + \Sigma_{k=1}^{K} x_{mt}^{k}(\sigma_k v_i^k + \pi_{k1}D_{i1} + \cdots + \pi_{kd}D_{id})\right]}, \tag{11}$$

- This integral is computed by simulation: doesn't necessarily require logit errors, though that is helpful and mixed logit is much more flexible than multinomial logit
- System of market share equations in full model solved numerically for δ_{jt} with contraction mapping:

$$\delta_{.t}^{h+1} = \delta_{.t}^{h} + lnS_{.t} - lnS(p_{.t}, x_{.t}, P_{ns}; \theta_2)$$

Estimation

Nevo (2001)

- In Nevo (2001) error term for GMM is city-quarter brand deviation from mean brand valuation. In BLP it is these two things combined
  - Introduces a challenge for β due to collinearity of attributes and mean brand effect
- GMM computed using a non-linear search, simplified by solving for linear parameters in terms of non-linear parameters and substituting
- Implementation of instruments:
  - Market specific markup correlated with market specific error term (compare to BLP)
  - BLP IVs of observed characteristics would be singular; not stellar anyway
  - Main approach: City-specific valuations are independent across cities controlling for brand mean and demographics

Estimation

Nevo (2001)

- Given this, price of products in other cities are valid IVs
  - Prices correlated due to common marginal cost
  - Uncorrelated with market specific valuation by assumption
- What do we think about this?
  - National or regional brand shocks
  - Regional advertising
- Examines additional IV of cost shifters, not a lot of variation however

Brand-Specific Dummies

Nevo (2001)

- Main reasons to include these:
  - Observed factors may not include everything, improves fit
  - Substitution patterns still driven by observed characteristics
  - Correlation between prices and unobserved quality is fully accounted for
- Two potential objections:
  - Dimensionality of parameters, not a major issue
  - Have to retrieve taste coefficients β
- Do this with minimum distance estimation:

$$d = X\beta + \xi$$

- Make same assumption as BLP IVs for this part $E[\xi|X] = 0$

Results

Nevo (2001)

- Use multinomial logit to show that IVs matter, then move on to full model
- Results for this part show:
  - Instruments have power, and impact coefficients of interest
  - Similarity of estimates from two sets of IVs
  - Importance of controlling for demographics
- Full model results presented for empirical CPS demographics, independent normal v, and logit ϵ
  - All coefficients statistically significant and of expected sign
  - Coefficients paint detailed picture of heterogeneity
  - Taste parameter standard deviations are insigificant BUT interaction with demographics IS significant in general E.G. Value of sogginess increases with age and income, older high income people tend to be less price sensitive

Full Model Results

Nevo (2001)

TABLE VI

RESULTS FROM THE FULL MODEL[a]

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			

[a] Based on 27,862 observations. Except where noted, parameters are GMM estimates. All regressions include brand and time dummy variables. Asymptotically robust standard errors are given in parentheses.
[b] Estimates from a minimum-distance procedure.

Full Model Results

Nevo (2001)

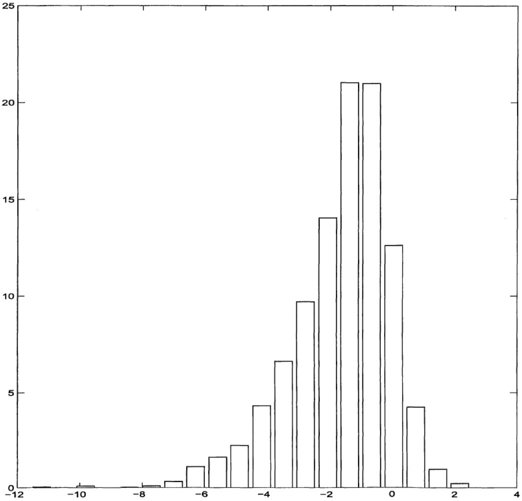

FIGURE 1.—Frequency distribution of taste for sogginess (based on Table VI).

Full Model Results

Nevo (2001)

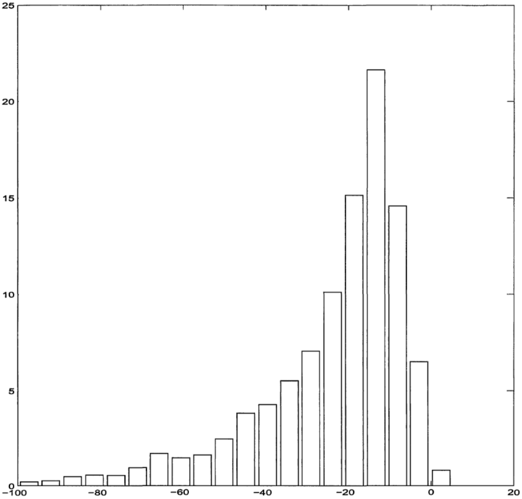

FIGURE 2.—Frequency distribution of price coefficient (based on Table VI).

Results

Nevo (2001)

- 90% of variation in heterogeneity comes from demographics and not unobserved parameters
- Full model result has important implications for cross-price elasticities:
  - Presents median of each entry over 1124 sample markets
  - Logit model restricts all elasticities within column in next table to be equal: shows how far this model comes from that model
  - Presents results on maximum to minimum ratio of cross price elasticities within a column
- Implications for Price-Cost Margins:
  - Markups different than logit, because of substitution patterns, see comparitive static with competition model
  - Match predictions to observed gross retail margins of 44%
  - Reject collusion, nothing else in full model

Full Model Results

Nevo (2001)

TABLE VII

MEDIAN OWN AND CROSS-PRICE ELASTICITIES[a]

#	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

[a] Cell entries i, j, where i indexes row and j column, give the percent change in market share of brand i with a one percent change in price of j. Each entry represents the median of the elasticities from the 1124 markets. The full matrix and 95% confidence intervals for the above numbers are available from http://elsa.berkeley.edu/~nevo.

Full Model Results

Nevo (2001)

TABLE VIII

MEDIAN MARGINS[a]

	Logit (Table V column ix)	Full Model (Table VI)
Single Product Firms	33.6%	35.8%
	(31.8%–35.6%)	(24.4%–46.4%)
Current Ownership of 25 Brands	35.8%	42.2%
	(33.9%–38.0%)	(29.1%–55.8%)
Joint Ownership of 25 Brands	41.9%	72.6%
	(39.7%–44.4%)	(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.

Discussion and Takeaways

Nevo (2001)

- Model improves burden on instruments from BLP, makes it more likely we are getting it right
- How much variation is there in pricing and demographics across regions to identify things? Structural setup and standard errors: how much does specification impact this?
- Do we believe conduct story in empirical application?
- Still using aggregate data but at least has explicit variation in markets, which is really valuable compared to BLP.
- Follow up work shows results potentially sensitive to starting parameter values
- A lot of assumptions going in here without a ton of data, have to break it down into pieces to do good analysis: trade-off between credibility and precision