BLP applications: Nevo (2001) and Petrin(2002)

February 16, 2009

1 Nevo 2001 (Econometrica): Measuring market power in the ready-to-eat cereal industry

Main question: The ready-to-eat (RTE) cereal industry is characterized by high price-to-cost margins (PCM) and high concentrations, and so has been accused of being an industry with collusive pricing behavior and intense non-price competition. Nevo examines this argument by estimating the PCM and decomposing it into 3 sources: 1) that due to product differentiation, 2) that due to multiproduct firm pricing and 3) that due to price collusion.

Basic idea: Use the BLP framework to estimate brand-level demand. Then use these demand estimates along with different pricing rules to back out PCMs. Compare these PCMs against crude measures of actual PCM to separate the different sources of the markup.

1.1 Demand Model

- Indirect utility is:

$$u_{ijt} = x_j\beta_i + \alpha_i p_{jt} + \xi_j + \Delta\xi_{jt} + \epsilon_{ijt} \tag{1}$$

- uses brand-level dummy variables (fixed effects) to capture the mean characteristics of RTE cereal. Because the fixed effect picks up all characteristics (observed and unobserved), what's "left" as a structural error is the change in ξ_j over time, or $\Delta\xi_{jt}$. Need to have a rich data set to use fixed effects.

- Also includes demographic variables – allows them to interact product characteristics (e.g. does a consumer have a child, interacted with if a cereal is sugary). Data are from the March Current Population Survey.

- e.g. $\alpha_i = \alpha +$ demographic component + demographic/random component

- Because has top 25 products over time, no change in characteristics or choice sets over time. So usual BLP instruments do not work.

- Paper has a nice section on instruments, which you should read. Nevo tries a couple of different instruments (all of which have their faults) to convince the reader of the robustness of the results.

1.2 Supply Model

- Suppose there are F firms each of which produces some subset of $\mathcal{F}_f$ of the $j = 1, \ldots, J$ brands in the RTE cereal industry. Profits are given by

$$\Pi_f = \sum_{j \in \mathcal{F}_j} (p_j - mc_j) M s_j(p) - C_f \tag{2}$$

where M is market size, s_j is market share, mc is marginal cost and C is fixed cost.

- Assuming a Bertrand-Nash equilibrium in prices, the price p_j of any product must satisfy the FOC

$$s_j(p) + \sum_{r \in \{_j} (p_r - mc_r) \frac{ds_r(p)}{dp_j} = 0 \tag{3}$$

- As before, define an ownership structure, where

$$\hat{\Omega}_{jr} = \begin{cases} 1, & \text{if } \exists f : \{r, j\} \subset \mathcal{F}_f \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

and

$$S_{jr} = -\frac{ds_r}{dp_j} \tag{5}$$

and so $\Omega_{jr} = \hat{\Omega}_{jr} * S_{jr}$.

- Get the usual markup equation

$$p - mc = \Omega^{-1} s(p) \tag{6}$$

- Unlike BLP which solved for both demand and supply side simultaneously, Nevo does it separately (and so the demand side estimates do not depend on the supply-side assumptions). The advantage of of imposing a supply side is that the extra structure typically gives you smaller standard errors. Nevo can forego this b/c of the quality of his data.

1.3 Data

- a market is a city-quarter.
- market shares and prices in each market (65 cities)
- brand characteristics (e.g. fiber contents, sugary), advertising and demographic information (last two are "extra")
- period: 1988:Q1 to 1992:Q4
- only uses 25 brands with the highest national market share
  - Why limit the subsample? Not sure.
  - In BLP (and others), main source of identification is the change in choice set over time. See how consumers react when products enter and exit. Here, there is no entry/exit. But do have lots of price variation.
  - Perhaps – the high amount of entry/exit of small products reflects dynamic forces (experimentation with characteristics, etc), that a static model would be bad at capturing.

1.4 Results

- Strategy is: estimate the demand side and compute the price elasticities. Then he can evaluate the PCM for three different industry conduct models.

  1. single-product firms – prices of each brand are set to maximize profits for that brand. Under this structure, see how much product differentiation drives PCM
  2. current industry structure (multiproduct firms) Under this structure, see how much product differentiation and multiproduct profit maximization drives PCM
  3. Joint profit maximization (i.e. collusion) Under this structure, see how collusive behavior and product differentiation drives PCM

- By comparing the PCM under the three structures can see which one matches the data best, as well as, decompose the sources of PCM.

- See table of price elasticities (Table VII)

- See table of median margins (Table VIII)

- In the data/literature: PCM estimates are 46%, lower bound of 31% and an upper bond of 64%. Conclusions: Collusion does not match the data (PCM would be much higher), but multiproduct pricing does. Large part of the PCM is driven by product differentiation.

- Final thought: There is a lot of entry and exit of products in cereal (not addressed here). Interesting to consider the endogeneity of product characteristics.

2 Petrin 2002 (Journal of Political Economy): Quantifying the Benefits of New Products: The Case of the Minivan

Main question: What are the consumer welfare benefits from the introduction of the minivan?

Basic idea: Augments the BLP framework with an extra set of moments (micro-moments), allows for more precise estimation of structural parameters.

2.1 Data

Data set one (basic)

- List price and sales of new vehicles in the US from 1981 to 1993 an annual basis
- Vehicle characteristics (e.g. weight, horsepower)

Data set two (extra) moments – from Consumer Expenditure Survey (CEX)

- demographics of purchasers – family size and age of head of household by vehicle type
- Nevo simply added demographics to indirect utility – Petrin incorporates the demographics in the indirect utility & adds extra moments.

2.2 Demand Side

- Indirect utility is:

$$u_{ij} = X_j\beta + \alpha_i \log(y_i - p_j) + \xi_j + \sum_k \gamma_k \nu_{ik} x_{jk} + \xi_j + \epsilon_{ijt} \tag{7}$$

where

$$\alpha_i = \begin{cases} \alpha_0, & \text{if } y_i \leq \bar{y}_1 \\ \alpha_1, & \text{if } \bar{y}_1 \leq y_i \leq \bar{y}_2 \\ \alpha_2, & \text{if } y_i \geq \bar{y}_3 \end{cases} \tag{8}$$

and where $(\bar{y}_1, \bar{y}_2)$ divide the US population into three equally sized groups.

note: in general it is easier to forgo the log specification and use $\frac{\alpha}{y_i}p_j$.

- ν are draws from $N(0,1)$, measuring deviations from mean value of a characteristic
- Make parametric assumptions about how demographics influence tastes of characteristics, where (mi, sw) stand for minivan and station wagon, and fv stands for family vehicle.

$$\gamma_{i,mi} = \gamma_{mi} \log(\text{familysize}_i)\nu_{i,fv} \tag{9}$$
$$\gamma_{i,sw} = \gamma_{sw} \log(\text{familysize}_i)\nu_{i,fv} \tag{10}$$
$$\tag{11}$$

where $\nu_{i,fv}$ is the iid shock for family vehicles. This parameterization implies that larger families will prefer minivans and station wagons (as seen in the data).

- Example of building the model to reflect the data (i.e. CEX). Don't want to build in the results, but also want to give the model a chance.

2.3 Supply Side

The usual framwork profits are:

$$\Pi_f = M \sum_{j \in \mathcal{F}_f} (p_j - mc_j)s_j \tag{12}$$

and you get easily derive the usual PCM equation.

2.4 Results

- Has the usual BLP moments $E[\xi_j|(X,W)] = E[\omega_j|(X,W)] = 0$ where ω_j are the structural errors on the supply side, and W are the observed components of cost.

- In addition, has 7 additional moments based on the data from the CEX. The first three pin down sensitivity to price

$$E[\{i \text{ purchases new vehicle}|\{y_i < \bar{y}_1\}], \tag{13}$$
$$E[\{i \text{ purchases new vehicle}|\{\bar{y}_1 < y_i < \bar{y}_2\}], \tag{14}$$
$$E[\{i \text{ purchases new vehicle}|\{y_i > \bar{y}_2\}] \tag{15}$$

and the last four deal with family vehicles

$$E[fs_i|i \text{ purchases a minivan}], \tag{16}$$
$$E[fs_i|i \text{ purchases a station wagon}], \tag{17}$$
$$E[fs_i|i \text{ purchases a SUV}], \tag{18}$$
$$E[fs_i|i \text{ purchases a fullsize van}], \tag{19}$$

- Add these additional 7 moments, by just stacking them with the BLP moments. Note the weighting matrix will be block-diagonal, since the two sets of moments come from independent sampling processes.

- Show table 4 in paper (highlight the α estimates)

2.5 Welfare analysis

- Compensating variation is used to measure changes in consumer welfare from the introduction of the minivan. This is the change in a consumer's income that equates utility in a particular economic environment to a choosen benchmark.

- Benchmark is the real world – with minivans, alternative environment one without.

- Counterfactual, use the FOC on the supply side to solve for prices *without* minivans.

- see Table 7 for some price changes when minivans disappear
- Compensating variation is the dollar amount needed to make consumers indifferent between counterfactual and real world. For purchases of minivans, how many dollars are needed to compensate them for buying something other than a minivan.
- see Table 8 for compensating variation. Note that OLS, IV, and RC all imply that minivan purchases dislike minivan characteristics relative to their second choice (see negative signs for obs char) which is implausible. Adding the micromoments fixes this. (Table 13 has compensating variation over all consumers).
- For change in profits, Petrin looks at markups in the observed case and in the counterfactual. See paper (table 11) for details.

General Comment

One needs to think carefully about how the model is designed and what is the main question to be answered. For example, if going after certain effects, adding moments can be key. Also, what you include in the model is a function of your question. For example, if going after welfare effects, it is crucial that the indirect utility inputs reflect stuff that you really believe could be in a person's utility function (e.g. horsepower). But if only going after price elasticities, that you can be more liberal and include things in the indirect utility that proxy for stuff you think are in the utility function.

Petrin spends a lot of time comparing results from the simpl logit to the full-blown random-coefficients. The theme is that the simplier models rely too much on the iid "logit" error. Putting in more flexibility with the observed characteristics reduces the model's dependence on the error term, and, in turn, leads to more sensible results.