FALL 2021

Problem Set 3: Estimating BLP

1 Nevo's Code

The first part of this problem set is from Aviv Nevo. It has two goals. First, it will introduce you to the use of a structural model of demand and supply in a differentiated good market for simulation of the effects of horizontal mergers. Second, it will lead you through the technical details of estimating Logit and Mixed Logit models using aggregate data. We strongly encourage you to work together and help each other in this problem set.

1. Download the data and use to it estimate the following (Logit) model:

$$u_{ijt} = \alpha p_{jt} + x_j\beta + \xi_j + \Delta\xi_{jt} + \epsilon_{ijt} \qquad i = 1, ..., I; j = 1, ..., J; t = 1, ..., T$$

where ϵ_{ijt} is distributed i.i.d. extreme value. The notation follows that used in class. The vector of characteristics x_j contains a constant, sugar content, and a mushy dumm variable (=1 if cereal gets soggy in milk). Note: the Logit model sets the "non-linear" part to zero, so you will only use what is called X_1 in the data (and not X_2).

Estimate the following 4 specifications: (i) OLS without brand fixed effects, (ii) OLS with brand fixed effects; (iii) IV (using the instrumental variables provided in the data) with and without brand fixed effects. For each specification report the estimated coefficients and their standard errors.

2. Using these results compute the markups predicted by a multi-product Nash-Bertrand equilibrium. What are the implied estimates of marginal costs? Explain how you computed these.Report the mean, median and standard deviation of the distribution of the markups, margins and implied marginal costs.

3. Use the pre-merger estimate of marginal costs, the estimated price elasticities and an assumption of multi-product firm Nash-Bertrand post-merger equilibrium to simulate the post merger equilibrium. Explain exactly each step. Simulate the effect of a Post-Nabisco merger and GM-Quaker merger. (A company can be identified by the first digit in the id variable. GM=2, Post=3, Quaker =4, Nabisco=6). Report the changes in the equilibrium prices and quantities.

4. Discuss the potential problems with the analysis you preformed in the previous question. How can you deal with these issues?

5. Obtain a computer code that estimates the Mixed Logit model with aggregate data using the NFXP algorithm we discussed in class (which is described in the "The Practitioner's Guide to ..."). You can either: (1) write the code yourself, (2) use the MATLAB code that we have provided or

(3) use the MATLAB code to help you program your own version. If you plan to write the code yourself see some guidance below. If you plan to take the code by Nevo, be sure you understand it.

Using this code and the above data, estimate the model

$$u_{ijt} = \alpha_i p_{jt} + x_j \beta_i + \xi_j + \Delta\xi_{jt} + \epsilon_{ijt} \qquad i = 1, ..., I; j = 1, ..., J; t = 1, ..., T$$

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i, \quad \nu_i \sim N(0, I)$$

where ϵ_{ijt} is distributed i.i.d extreme value. The notation follows that used in class. The vector of characteristics, x_j, contains a constant, sugar content and a mushy dummy (=1 if cereal gets soggy in milk). The demographic variables D_i include income, income squared, age and a child dummy (=1 if age $\leq$ 16). The constant and the coefficients on sugar and mushy should vary by income and age. The price coefficient should vary with income, income squared and child. Set all the off-diagonal elements of Σ to 0. Report the coefficients $(\alpha, \beta, \Pi, \Sigma)$ and their standard errors.

Hint: for the initial values of Σ and Π you might want to use something like (0.3, 5, 0, 0.2,0), (2.2, 13, -1, 0, 2.5), (0.01, -0.2, 0, 0, 03, 0), and (0.2, 1.3, 0, -0.8, 0) for the coefficients on the constant, price, sugar and mushy, respectively (for each coefficient the first number denotes σ and the next 4 the interactions with income, income squared, age and child. A 0 implies the coefficient is set to zero a-priori and should not be part of the optimization).

Be sure you get the same estimates that are reported in the **results** file.

6. Using these results repeat question 2. How did the markups change? Is this what you expected? Explain.

7. Using these new results repeat question 3. How did the effects of the mergers change? Is this what you expected? Explain.

To complete the problem set, the following papers are very helpful - we suggest you read them closely:

- Nevo, Aviv. "A practitioner's guide to estimation of random-coefficients logit models of demand." Journal of economics management strategy 9.4 (2000): 513-548.
- Nevo, Aviv. "Mergers with differentiated products: The case of the ready-to-eat cereal industry." The RAND Journal of Economics (2000): 395-421.
- Nevo, Aviv. "Measuring market power in the ready-to-eat cereal industry." Econometrica 69.2 (2001): 307-342.

2 PyBLP

In this part of the problem set, you are asked to familiarize yourself with PyBLP (Conlon and Gortmaker, 2020) and replicate results from the Nevo code. You can read about pyBLP at https://pyblp.readthedocs.io/en/stable/.

1. Estimate the same model as in 5. above. Do you obtain the same results? Why? Which settings seem to make a difference?

2. Obtain estimated markups and merger efects as in 6. and 7. above. How do the pyBLP results compare to your MATLAB results?

Report results and Python code as in Figure 5 in Conlon and Gortmaker (2020).

- Conlon, Christopher, and Jeff Gortmaker. "Best practices for differentiated products demand estimation with PyBLP." The RAND Journal of Economics 51.4 (2020): 1108-1161.