

Simulating the Effects of a Horizontal Merger

This problem set has two goals. First, it will introduce you to the use of the structural models we covered in class 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.

1. Download the data from the course web-page and use to it estimate the following (Logit) model:

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

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 dummy 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 algorithm we discussed in class (which is described in the “The Practitioner’s Guide to ...”). You can either: (1) write the code yourself, (2) download a MATLAB code from my home page; 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 my code be sure you understand it.

Using this code and the above data. Estimate the model

$$u_{ijt} = \alpha p_{jt} + x_j \beta_i + \xi_j + \Delta \xi_{jt} + \epsilon_{ijt} \quad i = 1, \dots, I, j = 1, \dots, J, t = 1, \dots, T,$$

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

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 include income, income squared, age and a child dummy (==1 if age<=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 (α , β , Π , and Σ) 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 no be part of the optimization).

Be sure you get the same estimates I report on my home page.

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

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

Guidance to programming the code yourself: Program Steps 1 through 4 given in Section 4 of “A Practitioner’s Guide to Random Coefficients Discrete Choice Models of Demand.” You can use any software you like. I recommend Matlab, but Gauss, C, or Fortran will also do. You should treat each of the steps as a separate (not too difficult) programming problem. You are NOT expected to write the most general program just one that can deal with the above data

The data is described in detail in question 2, but for now here is a description of the input of your program. You have data on T markets (47 cities * 2 quarters) for J (24) brands of cereal. The data contains a vector s_{jt} (94*24=2256 by 1) of market shares, a matrix $X1$ (2256 by 25) of variables that enter the mean utility level (prices+24 brand dummies), a matrix $X2$ (2256 by 4) of variables that enter the “non-linear” part, a matrix V (94 by 80) of random draws of v_i (each row is for a different market. The 1st 20 columns are draws for the 1st characteristic, 2nd 20 for the next characteristics and so on), and a matrix of demographic variables $DEMO$ (94 by 80). Each row is a market. The first 20 columns are different draws of income from the CPS for that market, next 20 are a different demographic variable for the same individuals, etc.). Finally, you will have a matrix IV of instrumental variables (2256 by 20).

The output of the program should be the estimated coefficients, both linear and non-linear (i.e. α , β , Π , and Σ), and their standard errors.

Here are some hints that might help you along the way. In Step 1 use the smooth simulator for market shares. This part of the program is the one used most often, so be sure to vectorize the computation as much as you can. When forming the GMM objective function use all the IV’s + plus the brand dummies, and a fixed weight matrix equal to $A = Z' * Z$ (where Z is the matrix of instruments). For Step 4 I recommend using either a Nelder-Mead simplex search method (fmins in Matlab), or a quasi-Newton method (fminu in Matlab) using the analytic gradient given in the Guide. The latter is much faster but requires additional programming. It has the additional advantage of helping in the computation of the standard errors.