

Problem Set 1: Empirical Demand Estimation and Merger Simulation

This problem set is designed to give you experience programming the kind of empirical demand and pricing models that we've discussed in weeks 2 and 3 of class. The material here is most closely related to BLP(1995) and Nevo (2001). Most of the material in this problem set was generated by Greg Crawford, Economics Professor at University of Warwick.

As stated in the syllabus, you may work in up to groups of three on this. This is mainly for you to have a commitment device to do the work yourself, since it is in your own best interest to do so. If you have to talk to another group for advice, it's no big deal.

To answer the questions here you will need a dataset (that has been generated by Crawford) that can be found on the course website. These data contain a cross-section of market shares, prices, and demand and cost shifters for five goods in each of 100 markets. Each good is offered in each market. This dataset mimics the kinds of aggregated market-level datasets studied in BLP and Nevo and frequently available in empirical industrial organization.

You can think of this data as being representative of a retail market in the U.S. (the car rental market). Given in the data are the following variables:

Varname	Variable
market	Market identifier, n (1-100)
prodid	Product (and firm) identifier, j (1-5)
sjn	Market share for j in n
pjn	Price for product j in n
d1-d5	Product (brand) dummy variables (demand shifters)
x1, x2	Product characteristics (demand shifters, e.g. # stores, quality rating)
w1, w2	Non-characteristic cost shifters (e.g. fuel cost, wages)

The problem set begins with a series of questions designed to think about key issues, such as the potential endogeneity of prices. We'll begin by doing some simple regressions, and move towards a final specification that is the full mixed logit model similar to that estimated in Nevo (2001). Most work can be done in STATA, but Matlab may be necessary for some of the more complex estimation in the latter questions.

1. Our first set of questions centers on a Logit model
 - (a) Estimate an aggregate Logit model using Ordinary Least Squares (OLS).
 - (You will have to create the dependent variable.)
 - For your OLS results, calculate the average (across markets) estimated own-price elasticities for each product and the average estimated cross-price elasticity with respect to product 3. Be sure to use the fitted value of the market share(s) when calculating your elasticities.
 - (b) Estimate the same Logit model using Instrumental Variables (IV) with cost-shifters as instruments.
 - Be sure to report the results of the first-stage regressions when you estimate by IV. Test whether the instruments are jointly significant.
 - (c) Estimate (again) the Logit model using IV with the average of the characteristics of other products as instruments.
 - (You will have to create this set of instruments from the raw data.)
 - As above, be sure to report the results of the first-stage regressions when you estimate by IV. Test whether the instruments are jointly significant.
 - (d) Which set of IV results do you prefer? Why? Call those your IV Logit results.
 - As above, calculate all the average estimated own-price and the average estimated cross-price elasticity w.r.t. product 3 for your IV Logit results
2. We now will move to the Nested Logit model, which we have not gone over explicitly in class but that you should be familiar with. This model is discussed in Nevo (2001) and NEvo (2000):
 - (a) Suppose you did some investigation into the car rental market and found that firms 2 and 3 both had a large presence at the airport, while firms 4 and 5 both had a large presence in the city center. Write down a tree structure to represent likely patterns of substitutability within the market given this fact.

- (b) Estimate an aggregate Nested Logit model reflecting this pattern of substitutability using OLS. Note for convenience you may assume the same nesting parameter (σ) captures the similarity between 2 and 3 as well as 4 and 5.¹
- (You will have to create the within-group share variable in order to do this.)
 - As above, calculate all the average estimated own-price and the average estimated cross-price elasticity w.r.t. product 3 for your OLS Nested Logit results.²
- (c) Estimate the same Nested Logit model using Instrumental Variables (IV) with cost-shifters as instruments
- Be sure to report the results of the first-stage regressions when you estimate by IV.³ Test whether the instruments are jointly significant in each case.
- (d) Estimate (again) the Logit model using IV with the (average of the) characteristics of other products *within the group* as instruments.
- (You will have to create this set of instruments from the raw data. Note the “average” x_1 for product 2 is just the x_1 for product 3 (and vice versa). There is no average other characteristic for product 1 as it is in its own group, so just give it a value of 0.)
 - As above, be sure to report the results of the first-stage regressions when you estimate by IV. Test whether the instruments are jointly significant.
- (e) Estimate (again) the Nested Logit model using IV with *both* cost-shifters and within-group average of rivals’ characteristics as instruments.
- As above, be sure to report the results of the first-stage regressions when you estimate by IV. Test whether the instruments are jointly significant.

¹One wouldn’t have to do this. Allowing them to differ would just mean we would have to estimate two nesting parameters instead of one.

²You need to calculate the fitted market shares (*and* within-group shares) in your nested logit model.

³Two notes: (1) Be sure to include all the other x ’s when you do your first-stage regressions, and (2) Note that you have *two* right-hand side endogenous variables now, p_{jn} and $\log(s_{jgn})$. Therefore you need to run first-stage regressions for *each*.

- (f) Which set of IV results do you prefer? Why? Call those your IV Nested Logit results.
- As above, calculate all the average estimated own-price and the average estimated cross-price elasticity w.r.t. product 3 for your IV Logit results
3. For each set of (four) sets of results above, calculate the marginal costs implied by your results for each product in each market. Which sets of marginal costs appear to be reasonable (there can be more than one)? Why?
- To help you answer this question, calculate the share of marginal costs across observations that are estimated to be negative. Why does this happen?
4. Suppose you were working for the competition authority when firms 2 and 3 announce their intention to merge.
- Derive what the pricing equation is for good 2 and good 3 under the assumption of logit tastes when firms 2 and 3 merge.⁴ Using this equation, predict the average percentage change in prices for goods 2 and 3 using your IV Logit results.
 - Derive what the pricing equation is for good 2 and good 3 under the assumption of nested logit tastes when firms 2 and 3 merge.⁵ Predict the average percentage change in prices for goods 2 and 3 using your IV Nested Logit results.
 - Are they qualitatively different? If so, provide an economic/econometric explanation for why this is so. Which do you think more plausible? Why?
5. Suppose I told you that the data weren't generated by a Nested Logit, but by a Random Coefficients Model (RCM), which is true. To answer this question you will want to understand and modify code written by Nevo on how to estimate an RCM. This code can be found here:

⁴Focus just on goods 2 and 3 - we will assume prices for goods 1, 4, and 5 stay the same. To figure out the pricing equation, you will have to plug in the values for S_{jr} for the logit and then invert the 2×2 matrix given by the part of Ω corresponding to goods 2 and 3 to answer this part.

⁵Here you do just as you did above, except that the elements in S_{jr} are more complicated.

http://faculty.wcas.northwestern.edu/~ane686/supplements/rc_dc_code.htm.

Use this to answer the following questions:

- (a) Estimate the parameters of the RCM using the same instruments you used for your preferred IV Nested Logit results. You will need to do this in Matlab or Gauss (I did it in Matlab). Nevo's code is a good place to start (although it will need to be modified in some ways).
- (b) Based on your coefficient estimates, simulate the effects of a merger between firms 2 and 3. By how much do prices for goods 2 and 3 increase? How does that compare with your estimates from your Logit and Nested Logit models?
- (c) Let's bring the supply-side into the estimation. Jointly estimate the demand and supply sides of a simple logit model. This means that there is only a single parameter that enters both: α . Be sure to include standard errors!
- (d) Jointly estimate the demand and supply sides of the random-coefficient model.
- (e) (Optional, but valuable) Construct your own monte carlo exercise of a differentiated product market. Create your own data. (Note: to do so, you will need to construct demand and cost shifters, error terms, and solve for equilibrium prices in each market.⁶) I would recommend you create the data from a RCM (as that is most flexible). Using the same code that you used to answer the remaining questions in this problem set, estimate the structural parameters you chose in your monte carlo. How well does it do?

⁶That is analogous to what you did when solving for the prices after a merger.