Industrial Organization II (ECO 2901)

Winter 2012. Victor Aguirregabiria

Problem Set #1

Demand; Static Models of Bertrand Competition; Exogenous Mergers

Due on Thursday, March 1st, 2012

TOTAL NUMBER OF POINTS: 100

In this problem set we propose a model of competition in a differentiated product industry, study some properties of the model, estimate its structural parameters using actual data, and use the estimated model to predict the effects of a merger.

There are F firms competing in this industry. We index firms by $f \in \{1, 2, ..., F\}$. These firms sell a total of J products, and we index products by j. The set of all products is $\mathcal{J} = \{1, 2, ..., J\}$, and the set products sold by firm f is $\mathcal{J}_f$ that is a subset of $\mathcal{J}$. The outside product (no purchase) is represented by the index $j = 0$. The profit of firm f is:

$$\pi_f = \sum_{j \in \mathcal{J}_f} [p_j \; q_j - c_j(q_j)]$$

with the obvious definitions. Consumer demand is characterized by a discrete choice model, and more specifically by a Nested Logit model (Ben-Akiva, 1973). A consumer indirect utility of buying product j is:

$$V_j = \delta_j + \varepsilon_j = X_j \; \beta - \alpha \; p_j + \xi_j + \varepsilon_j$$

with the definitions that you know. Consumer taste heterogeneity is captured by the term ε_j that in the Nested Logit model has the following structure: $\varepsilon_j = \sigma \; \varepsilon_{g_j}^{(1)} + \varepsilon_j^{(2)}$, where σ is a positive parameter, and $\varepsilon_{g_j}^{(1)}$ and $\varepsilon_j^{(2)}$ are independent variables with an Extreme Value type 1 distribution. g_j represents the group of product j in a partition of the set of products $\mathcal{J}$ into G groups. The idea is that products within the same group share common features that make them closer substitutes than products in different groups. Let $s_j \equiv q_j / H$ be the market share of product j, where H is the number of consumers in the market. Given this specification of utility, consumer optimal behavior implies the following equation for market shares:

$$s_j = s_{g_j}^* \; s_{j|g_j}$$

$s_{g_j}^*$ represents the share of consumers who choose a product within group g_j. Share $s_{j|g_j}$ represents the proportion of consumers who choose product j within the subset of consumers who select group g_j. The outside alternative, 0, is treated as a separate group with only one choice alternative, i.e., group 0 with $j = 0$. These market shares have the following form:

$$s_{j|g_j} = \frac{\exp\{\delta_j\}}{\sum_{k \in g_j} \exp\{\delta_k\}}$$

and

$$s^*_{g_j} = \frac{\exp\left\{\frac{IV_{g_j}}{\sigma}\right\}}{\sum_{g=0}^{G} \exp\left\{\frac{IV_g}{\sigma}\right\}}$$

where $\{IV_g : g = 0, 1, ..., G\}$ are the *inclusive values* that are defined as:

$$IV_g \equiv \mathbb{E}\left(\max_{j \in g}\left\{\delta_j + \varepsilon_j^{(2)}\right\} \mid \delta\right) = \ln\left(\sum_{j \in g} \exp\{\delta_j\}\right) \quad \text{for } g > 0$$

and $IV_0 = 0$.

Question 1 [5 points]: Show that the Nested Logit models implies the following system of equations relating market shares and average utilities ($\delta' s$).

$$\ln\left(s_{j|g_j}\right) = \delta_j - \sigma \ \ln\left(\frac{s^*_{g_j}}{s^*_0}\right)$$

Question 2 [5 points]: Suppose that you have a dataset where you observe $\{q_{jm}, p_{jm}, H_m, X_{jm}\}$ for every product j in the industry and over M local markets indexed by m. The number of local of markets M is relatively large (e.g., 2,000 markets) and the number of products is relatively small (e.g., 20 products). Describe an approach to estimate consistently the demand parameters β, α, and σ taking into account that prices p_{jm} can be correlated with unobservables ξ_{jm}.

Question 3 [5 points]: Suppose that the number of consumers in a market, or market size, H_m is measured with error. We observe H_m but the true market size is H_m^{true}, such that $H_m = H_m^{true} + e_m$ and e_m is measurement error. Propose a simple method to deal with this measurement error that does not require any specific assumption about the distribution of the error. [Hint: It is possible to show that this measurement error enters additively in our regression equations and is the same for every product j.]

Question 4 [5 points]: Derive close-form expressions for the following partial derivatives in the Nested Logit demand system: $\frac{\partial q_j}{\partial p_j}$; $\frac{\partial q_k}{\partial p_j}$ for $k \neq j$ and k, j in the same group g; and $\frac{\partial q_k}{\partial p_j}$ for k and j in different groups.

Question 5 [5 points]: Suppose that firms in this industry compete in prices ala Bertrand-Nash. Using the expressions that you have derived in Question 4, obtain the expression for the best response pricing equations of a firm: (a) when each firm produces a single product; and (b) when firms produce multiple products.

The STATA datafile `eco2901_problemset_01_2012_airlines_data.dta` contains a panel dataset similar to the one described in Question 2. It contains data of the US airline industry in 2004. A market is a *route* or directional city-pair, e.g., round-trip Boston to Chicago. A product is the combination of route (m), airline (f), and the indicator of stop flight or nonstop flight. For instance, a round-trip Boston to Chicago, non-stop, with American Airlines is an example of product.

Products compete with each other at the market (route) level. Therefore, the set of products in market m consists of all the airlines with service in that route either with nonstop or with stop flights. The dataset contains $2,950$ routes, 4 quarters, and 11 airlines (where the airline "Others" is a combination of multiple small airlines). The following table includes the list of variables in the dataset and a brief description.

Variable name		**Description**
`route_city`	:	Route: Origin city to Destination City
`route_id`	:	Route: Identification number
`airline`	:	Airline: Name (Code)
`direct`	:	Dummy of Non-stop flights
`quarter`	:	Quarter of year 2004
`pop04_origin`	:	Population Origin city, 2004 (in thousands)
`pop04_dest`	:	Population Destination city, 2004 (in thousands)
`price`	:	Average price: route, airline, stop/nonstop, quarter (in dollars)
`passengers`	:	Number of passengers: route, airline, stop/nonstop, quarter
`avg_miles`	:	Average miles flown for route, airline, stop/nonstop, quarter
`HUB_origin`	:	Hub size of airline at origin (in million passengers)
`HUB_dest`	:	Hub size of airline at destination (in million passengers)

In all the models of demand that we estimate below, we include time-dummies and the following vector of product characteristics:

{ `price, direct dummy, avg_miles, HUB_origin, HUB_dest, airline dummies` }

In some estimations we also include market (route) fixed effects. For the construction of market shares, we use as measure of market size (total number of consumers) the average population in the origin and destination cities, in number of people, i.e., `1000*(pop04_origin + pop04_dest)/2`.

Question 6 [15 points]: Estimate a Standard Logit model of demand: (a) by OLS without route fixed effects; (b) by OLS with route fixed effects. Interpret the results. What is the average consumer willingness to pay (in dollars) for a nonstop flight (relative to a stop flight), ceteris paribus? What is the average consumer willingness to pay for one million more people of hub size in the origin airport, ceteris paribus? What is the average consumer willingness to pay for Continental relative to American Airlines, ceteris paribus? Based on the estimated model, obtain the average elasticity of demand for Southwest products. Compare it with the average elasticity of demand for American Airline products.

Question 7 [15 points]: Consider a Nested Logit model where the first nest consists of the choice between groups "Stop", "Nonstop", and "Outside alternative", and the second nest consists in the choice of airline. Estimate this Nested Logit model of demand: (a) by OLS without route fixed effects; (b) by OLS with route fixed effects. Interpret the results. Answer the same questions as in Question 6.

Question 8 [15 points]: Consider the Nested Logit model in Question 7. Propose and implement an IV estimator that deals with the potential endogeneity of prices. Justify your choice of

instruments, e.g., BLP, or Hausman-Nevo, or Arellano-Bond, ... Interpret the results. Compare them with the ones from Question 7.

Question 9 [15 points]: Given your favorite estimation of the demand system, calculate price-cost margins for every observation in the sample. Use these price cost margins to estimate a marginal cost function in terms of all the product characteristics, except price. Assume constant marginal costs. Include also route fixed effects. Interpret the results.

Question 10 [15 points]: Consider the route Boston to San Francisco ("BOS to SFO") in the fourth quarter of 2004. There are 13 active products in this route-quarter, from which 5 are non-stop products. The number of active airlines is 8: with both stop and non-stop flights, America West (HP), American Airlines (AA), Continental (CO), US Airways (US), and United (UA); and with only stop flights, Delta (DL), Northwest (NW), and "Others". Consider the "hypothetical" (in 2004) merger between Delta and Northwest. The new airline, say DL-NW, has airline fixed effects, in demand and costs, equal to the average of the fixed effects of the merging companies DL and NW. As for the characteristics of the new airline in this route: `avg_miles` is equal to the minimum of `avg_miles` of the two merging companies; `HUB_origin = 45`; `HUB_dest = 36`; and the new airline still only provides stop flights in this route.

(a) Using the estimated model, obtain airlines profits in this route-quarter before the hypothetical merger.

(b) Calculate equilibrium prices, number of passengers, and profits , in this route-quarter after the merger. Comment the results.

(c) Suppose that, as the result of the merger, the new airline decides also to operate non-stop flights in this route. Calculate equilibrium prices, number of passengers, and profits , in this route-quarter after the merger. Comment the results.