Industrial Organization II (ECO 2901)

Winter 2014. Victor Aguirregabiria

Selection of Problems and Questions

Outline

Problem set 2008

Industrial Organization II

Spring 2008. Victor Aguirregabiria

PROBLEM SET:
ESTIMATION OF STATIC AND DYNAMIC GAMES OF MARKET ENTRY
DUE ON THURSDAY, APRIL 3, 2008

This problem set describes a dynamic game of entry/exit in an oligopoly market. To answer the questions below, you have to write computer code (e.g., GAUSS, MATLAB) for the solution, simulation and estimation of the model. Please, submit the program code together with your answers.

Consider the UK fast food industry during the period 1991-1995, as analyzed by Toivanen and Waterson (RAND, 2005). During this period, the industry was dominated by two large retail chains: McDonalds (MD) and Burger King (BK). The industry can be divided into isolated/independent local markets. Toivanen and Waterson consider local districts as the definition of local market (of which there are almost 500 in UK). At each local market these retail chains decide whether to have an outlet or not.

We index firms by $i \in \{MD, BK\}$ and time (years) by t. The current profit of firm i in a local market is equal to variable profits, VP_{it}, minus fixed costs of operating an outlet, FC_{it}, and minus the entry cost of setting up an outlet by first time, EC_{it}. Variable profits are $VP_{it} = (p_{it} - c_i)q_{it}$, where p_{it} represents the price, c_i is firm i's marginal cost (i.e., the marginal cost of an average meal in chain i), and q_{it} is the quantity sold (i.e., total number of meals served in the outlet at year t). The demand for an outlet of firm i in the local market is:

$$q_{it} = \frac{S_t \ \exp\{w_i - \alpha \ p_{it}\}}{1 + \exp\{w_i - \alpha \ p_{it}\} + a_{jt} \ \exp\{w_j - \alpha \ p_{jt}\}}$$

S_t represents the size of the local market at period t (i.e., total number of restaurant meals over the year). w_i and w_j are the average willingness to pay for products i and j, respectively. α is a parameter. And a_{jt} is the indicator of the event "firm j is active in the local market at period t". Every period t, the active firms compete in prices. There is not dynamics in consumers demand or in variable costs, and therefore price competition is static. Fixed costs and entry costs have the following form:

$$\begin{aligned} FC_{it} &= FC_i + \varepsilon_{it} \\ EC_{it} &= (1 - a_{i,t-1}) \ EC_i \end{aligned}$$

The fixed cost is paid every year that the firm is active in the market. The entry cost, or setup cost, is paid only if the firm was not active at previous year (if $a_{i,t-1} = 0$). Both fixed costs and entry costs are firm-specific. The entry cost is time invariant. ε_{it} represents a firm-idiosyncratic shock in firm i's fixed cost that is iid over firms and over time with a distribution $N(0, \sigma^2)$. We also assume that ε_{it} is private information of firm i. If a firm is not active in the market, its profit is zero. For notational simplicity I "normalize" the variance of ε_{it} to be 1, though it should be understood that the structural parameters in the profit function are identified up to scale.

QUESTION 1. [5 POINTS] Consider the static model of price competition. Show that equilibrium price-cost margins, $p_{it} - c_i$, and equilibrium market shares, q_{it}/S_t, do not depend on market size S_t. Therefore, we can write the equilibrium variable profit function as:

$$VP_{it} = (1 - a_{jt})\ S_t\ \theta_i^M + a_{jt}\ S_t\ \theta_i^D$$

where θ_i^M and θ_i^D represent the equilibrium variable profits per-capita (per-meal) when firm i is a monopolist and when it is a duopolist, respectively.

The payoff-relevant information of firm i at period t is $\{x_t, \varepsilon_{it}\}$ where $x_t \equiv \{S_t, a_{1,t-1}, a_{2,t-1}\}$. Let $P_j(x_t)$ represents firm $i's$ belief about the probability that firm j will be active in the market given state x_t. Given this belief, the expected profit of firm i at period t is:

$$\begin{aligned} \pi_{it}^P &= (1 - P_j(x_t))\ S_t\ \theta_i^M + P_j(x_t)\ S_t\ \theta_i^D - FC_i - (1 - a_{i,t-1})\ EC_i - \varepsilon_{it} \\ &= Z_{it}^P\ \theta_i - \varepsilon_{it} \end{aligned}$$

where $Z_{it}^P \equiv ((1 - P_j(x_t))\, S_t, P_j(x_t)S_t, -1, -(1 - a_{i,t-1}))$ and $\theta_i \equiv \left(\theta_i^M, \theta_i^D, FC_i, EC_i\right)'$.

For the rest of this problem set, we consider the following values for the profit parameters:

$$\theta_{MD}^M = 1.5 \quad ; \quad \theta_{MD}^D = 0.7 \quad ; \quad FC_{MD} = 6 \quad ; \quad EC_{MD} = 6$$

$$\theta_{BK}^M = 1.2 \quad ; \quad \theta_{BK}^D = 0.3 \quad ; \quad FC_{BK} = 4 \quad ; \quad EC_{BK} = 4$$

MD's product has higher quality (even after adjusting for marginal costs) than BK's. This implies that MD has higher variable profits than BK, either under monopoly or under duopoly. However, MD has also higher costs of setting up and operating an outlet.

Market size S_t follows a discrete Markov process with support $\{4, 5, 6, 7, 8, 9\}$ and transition probability matrix:

$$F_S = \begin{bmatrix} 0.9 & 0.1 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.1 & 0.8 & 0.1 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.1 & 0.8 & 0.1 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.1 & 0.8 & 0.1 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.1 & 0.8 & 0.1 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.1 & 0.9 \end{bmatrix}$$

A. STATIC (MYOPIC) ENTRY-EXIT GAME

We first consider a static (not forward-looking) version of the entry-exit game. A Bayesian Nash Equilibrium (BNE) in this game can be described as a pair of probabilities, $\{P_{MD}(x_t), P_{BK}(x_t)\}$ solving the following system of equations:

$$\begin{aligned} P_{MD}(x_t) &= \Phi\left(Z_{MDt}^P\, \theta_{MD}\right) \\ P_{BK}(x_t) &= \Phi\left(Z_{BKt}^P\, \theta_{BK}\right) \end{aligned}$$

where $\Phi(.)$ is the CDF of the standard normal.

QUESTION 2. [10 POINTS] For every possible value of the state x_t (i.e., 24 values) obtain all the BNE of the static entry game.

Hint: Define the functions $f_{MD}(P) \equiv \Phi\left(Z_{MDt}^{P}\ \theta_{MD}\right)$ and $f_{BK}(P) \equiv \Phi\left(Z_{BKt}^{P}\ \theta_{BK}\right)$. Define also the function $g(P) \equiv P - f_{MD}(f_{BK}(P))$. A BNE is zero of the function $g(P)$. You can search for all the zeroes of $g(P)$ in different ways, but in this case the simpler method is to consider a discrete grid for P in the interval $[0,1]$, e.g., uniform grid with 101 points.

For some values of the state vector x_t, the static model has multiple equilibria. To answer Questions 3 to 5, assume that, in the population under study, the "equilibrium selection mechanism" always selects the equilibrium with the higher probability that MD is active in the market.

Let X be the set of possible values of x_t. And let $\mathbf{P}^0 \equiv \{P_{MD}^0(x), P_{BK}^0(x) : x \in X\}$ be the equilibrium probabilities in the population. Given $\mathbf{P}^0$ and the transition probability matrix for market size, F_S. We can obtain the steady-state distribution of x_t. Let $p^*(x_t)$ be the steady-state distribution. By definition, for any $x_{t+1} \in X$:

$$\begin{aligned} p^*(x_{t+1}) &= \sum\nolimits_{x_t \in X} p^*(x_t)\ \Pr\left(x_{t+1}|x_t\right) \\ &= \sum\nolimits_{x_t \in X} p^*(x_t)\ F_S(S_{t+1}|S_t) \\ & \quad \left[P_{MD}^0(x_t)\right]^{a_{MDt+1}} \left[1 - P_{MD}^0(x_t)\right]^{1-a_{MDt+1}} \left[P_{BK}^0(x_t)\right]^{a_{BKt+1}} \left[1 - P_{BK}^0(x_t)\right]^{1-a_{BKt+1}} \end{aligned}$$

QUESTION 3. [10 POINTS] Compute the steady-state distribution of x_t in the population.

QUESTION 4. [20 POINTS] Using the values of P^0, F_S and p^* obtained above, simulate a data set $\{x_{mt} : t = 0, 1, ..., T; m = 1, 2, ..., M\}$ for $M = 500$ local markets and $T + 1 = 6$ years with the following features: (1) local markets are independent; and (2) the initial states x_{m0} are random draws from the steady-state distribution p^*. Present a table with the mean values of the state variables in x_t and with the sample frequencies for the following events: (1) MD is a monopolist; (2) BK is a monopolist; (3) duopoly; (4) MD is active given that (conditional) he was a monopolist at the beginning of the year (the same for BK); (5) MD is active given that BK was a monopolist at the beginning of the year (the same for BK); (6) MD is active given that there was a duopoly at the beginning of the year (the same for BK); and (7) MD is active given that there were no firms active at the beginning of the year (the same for BK).

QUESTION 5. [20 POINTS] Use the simulated data in Question 4 to estimate the structural parameters of the model. Implement the following estimators: (1) two-step PML using a frequency estimator of $\mathbf{P}^0$ in the first step; (2) two-step PML using random draws from a

U(0,1) for $\mathbf{P}^0$ in the first step; (3) 20-step PML using a frequency estimator of $\mathbf{P}^0$ in the first step; (4) 20-step PML using random draws from a U(0,1) for P^0 in the first step; and (5) NPL estimator based on 10 NPL fixed points (i.e., 10 different initial $P's$). Comment the results.

QUESTION 6. [30 POINTS] Suppose that the researcher knows that local markets are heterogeneous in their market size, but he does not observed market size S_{mt}. Suppose that the researcher assumes that market size is constant over time but it varies across markets, and it has a uniform distribution with discrete support $\{4, 5, 6, 7, 8, 9\}$. Obtain the NPL estimator under this assumption (use 20 NPL fixed points). Comment the results.

QUESTION 7. [30 POINTS] Use the previous model (both the true model and the model estimated in Question 5) to evaluate the effects of a value added tax. The value added tax is paid by the retailer and it is such that the parameters θ_i^M and θ_i^D are reduced by 10%. Obtain the effects of this tax on average firms' profits, and on the probability distribution of market structure.

B. DYNAMIC ENTRY-EXIT GAME

Now, consider the dynamic (forward-looking) version of the entry-exit game. A Markov Perfect Equilibrium (MPE) in this game can be described as a vector of probabilities $\mathbf{P} \equiv \{P_i(x_t) : i \in \{MD, BK\}, x_t \in X\}$ such that, for every (i, x_t):

$$P_i(x_t) = \Phi\left(\tilde{Z}_{it}^P \; \theta_{MD} + \tilde{e}_{it}^P\right)$$

where $\tilde{Z}_{it}^P$ and $\tilde{e}_{it}^P$ are defined in the class notes.

QUESTION 8. [20 POINTS] Obtain the MPE that we obtain when we iterate in the equilibrium mapping starting with an initial $\mathbf{P} = 0$. Find other MPEs.

QUESTION 9. [10 POINTS] Compute the steady-state distribution of x_t in the population.

QUESTION 10. [20 POINTS] The same as in Question 4 but using the dynamic game and the MPE in Question 8.

QUESTION 11. [20 POINTS] The same as in Question 5 but using the dynamic game and the MPE in Question 8.

QUESTION 12. [30 POINTS] The same as in Question 6 but using the dynamic game and the MPE in Question 8.

QUESTION 13. [30 POINTS] The same as in Question 7 but using the dynamic game and the MPE in Question 8.

Final exam 2008

ECO 2901S

Industrial Organization II

Spring 2008. Victor Aguirregabiria

TEST (Due on Monday, April 14 at noon)

QUESTION 1 (25 POINTS): This question deals with the paper by Hendel and Nevo (*Econometrica,* 2006).

(a) Explain the implications on estimated elasticities and market power of ignoring (when present) consumer forward-looking behavior and dynamics in the demand of differentiated storable products. Discuss how the biases depend on the stochastic process of prices (e.g., Hi-Lo pricing versus a more stable price).

(b) Describe the main issues in the estimation of Hendel-Nevo model. Discuss the assumptions that they make to deal with these issues.

QUESTION 2 (25 POINTS): The geographic definition of a local market is an important modelling decision in empirical models of market entry.

(a) Explain the implications on the empirical predictions of these model of using a definition of local that is too broad or too narrow.

(b) Explain the approach in Seim (2006). Discuss its advantages and limitations.

QUESTION 3 (50 POINTS): There is a significant number of empirical applications of static and dynamic models of entry in local markets which find the following empirical regularity: after conditioning on observable market characteristics (e.g., population, income, age) there is a positive correlation between the entry decisions of potential entrants. Three main hypotheses have been proposed to explain this evidence: (1) spillover effects in consumer traffic; (2) information externalities (see Caplin and Leahy [*Economic Journal,* 1998] and Toivanen and Waterson [*RAND*, 2005]); and (3) market characteristics which are observable for the firms but unobservable to the researcher.

(a) Explain how these hypotheses can explain the empirical evidence.

(b) Discuss why it is important to distinguish between these hypothesis. Do they have different policy implications?

(c) Consider the data and the empirical application in Toivanen and Waterson (*RAND*, 2005). Explain how it is possible to identify empirically the contribution of the three hypotheses.

(d) Consider the dynamic game of entry-exit in the Problem Set of this course. Explain how to extend this model to incorporate information externalities as in Caplin and Leahy (1998). Discuss identification issues.

Problem set 2010

Industrial Organization II (ECO 2901)
Spring 2010. Victor Aguirregabiria

Problem Set #1. Static Entry Models
Due on Thursday, February 11, 2010

The Stata datafile `eco2901_problemset_01_chiledata_2010.dta` contains a panel dataset of 167 local markets in Chile with annual information over the years 1994 to 1999 and for five retail industries: Restaurants ('Restaurantes,' product code 63111); Gas stations ('Gasolineras,' product code 62531); Bookstores ('Librerias,' product code 62547); Shoe Shops ('Calzado,' product code 62411); and Fish shops ('Pescaderias,' product code 62141). The 167 "isolated" local markets in this dataset have been selected following criteria similar to the ones in Bresnahan and Reiss (1991). This is the list of variables in the dataset with a brief description of each variable:

`comuna_code`	:	Coder of local market
`comuna_name`	:	Name of local market
`year`	:	Year
`procode`	:	Code of product/industry
`proname`	:	Name of product/industry
`pop`	:	Population of local market (in # people)
`areakm2`	:	Area of local market (in square Km)
`expc`	:	Annual expenditure per capita in all retail products in the local market
`nfirm`	:	Number of firms in local market and industry at current year
`nfirm_1`	:	Number of firms in local market and industry at previous year
`entries`	:	Number of new entrants in local market and industry during current year
`exits`	:	Number of exiting firms in local market and industry during current year

Consider the following static entry model in the spirit of Bresnahan and Reiss (JPE, 1991, hereinafter *BR-91*). The profit of an active firm in market m at year t is:

$$\Pi_{mt} = S_{mt} \ v(n_{mt}) - F_{mt}$$

where S_{mt} is a measure of market size; n_{mt} is the number of firms active in the market; $v(.)$ is the *variable profit per capita* and it is a decreasing function; and F_{mt} represents fixed operating costs in market m at period t. The function $v(.)$ is nonparametrically specific. The specification of market size is:

$$S_{mt} = POP_{mt} \ \exp\left\{\beta_0^S + \beta_1^S \ expc_{mt} + \varepsilon_{mt}^S\right\}$$

where POP_{mt} is the population in the local market; $expc_{mt}$ is per capita sales in all retail industries operating in the local market; β_0^S and β_1^S are parameters; and ε_{mt}^S is an unobservable component of market size. The specification of the fixed cost is:

$$F_{mt} = \exp\left\{\beta^F + \varepsilon_{mt}^F\right\}$$

where β^F is a parameter, and ε^F_{mt} in an unobservable component of the fixed cost. Define the unobservable $\varepsilon_{mt} \equiv \varepsilon^S_{mt} - \varepsilon^F_{mt}$. And let $X_{mt} \equiv (\ln POP_{mt}, expc_{mt})$ be the vector with the observable characteristics of the local market. We assume that ε_{mt} is independent of X_{mt} and iid over (m,t) $N(0,\sigma^2)$.

Question 1. **[10 points]** Show that the model implies the following probability distribution for the equilibrium number of firms: let $n_{\max}$ be the maximum value of n_{mt}, then for any $n \in \{0, 1, ..., n_{\max}\}$:

$$\begin{aligned}
\Pr\left(n_{mt} = n \mid X_{mt}\right) &= \Pr\left(cut(n) \leq X_{mt} \begin{bmatrix} \frac{1}{\sigma} \\ \frac{\beta_1^S}{\sigma} \end{bmatrix} + \frac{\varepsilon_{mt}}{\sigma} \leq cut(n+1) \right) \\
&= \Phi\left(cut(n+1) - X_{mt} \begin{bmatrix} \frac{1}{\sigma} \\ \frac{\beta_1^S}{\sigma} \end{bmatrix} \right) - \Phi\left(cut(n) - X_{mt} \begin{bmatrix} \frac{1}{\sigma} \\ \frac{\beta_1^S}{\sigma} \end{bmatrix} \right)
\end{aligned}$$

where $cut(0)$, $cut(1)$, $cut(2)$, ... are parameters such that for $n \in \{1, 2, ..., n_{\max}\}$, $cut(n) \equiv (\beta^F - \beta^S_0 - \ln v(n))/\sigma$, and $cut(0) \equiv -\infty$, and $cut(n_{\max}+1) \equiv -\infty$.

Question 2. **[20 points]** Given the Ordered Probit structure of the model, estimate the vector of parameters $\{1/\sigma,\ \beta^S_1/\sigma,\ cut(1),\ cut(2),\ ...,\ cut(n_{\max})\}$ for each of the five industries separately. Given these estimates, obtain estimates of the parameters $\dfrac{v(n+1)}{v(n)}$ for $n \in \{1, 2, ..., n_{\max}\}$. Present a figure of the estimated function $\dfrac{v(n+1)}{v(n)}$ for each of the five industries. Interpret the results. Based on these results, what can we say about *the nature of competition* in each of these industries?

Question 3. **[20 points]** Repeat the same exercise as in Question 3 but using the following specification of the unobservable ε_{mt}:

$$\varepsilon_{mt} = \gamma_t + \delta_m + u_{mt}$$

where γ_t are time effects that can be captured by using time-dummies; δ_m are fixed market effects that can be captured by using market-dummies; and u_{mt} is independent of X_{mt} and iid over (m,t) $N(0,\sigma^2)$. Comment the results.

Now, consider the following static entry model of incomplete information. There are N_{mt} potential entrants in market m at period t. The profit of an active firm in market m at year t is:

$$\Pi_{imt} = S_{mt}\ v(n_{mt}) - F_{imt}$$

Market size, S_{mt}, has the same specification as in Question 2. The firm-specific fixed cost, F_{mt}, has the following specification:

$$F_{imt} = \exp\left\{ \beta^F + \varepsilon^F_{mt} + \xi_{imt} \right\}$$

The random variables ε^S_{mt}, ε^F_{mt}, and ξ_{imt} are unobservable to the researcher. From the point of view of the firms in the market, the variables ε^S_{mt} and ε^F_{mt} are common knowledge, while ξ_{imt} is private information of firm i. We assume that ξ_{imt} is independent of X_{mt} and iid over (m,t) $N(0,\sigma^2_\xi)$.

The number of potential entrants, N_{mt}, is assumed to be proportional to population: $N_{mt} = \lambda \; POP_{mt}$, where the parameter λ is industry specific.

Question 4. [5 points] Consider the following estimator of the number of potential entrants:

$$\hat{N}_{mt} = \text{integer}\left\{ \max_{\text{over all}\{m',t'\}} \left[\frac{entrants_{m't'} + incumbents_{m't'}}{POP_{m't'}} \right] \; POP_{mt} \right\}$$

where $entrants_{m't'}$ and $incumbents_{m't'}$ are the number of new entrants and the number of incumbents, respectively, in market m' at period t'. Show that $\hat{N}_{mt}$ is a consistent estimator of $N_{mt} = \lambda \; POP_{mt}$.

Question 5. [15 points] Let $P(X_{mt}, \varepsilon_{mt})$ be the equilibrium probability of entry given the common knowledge variables $(X_{mt}, \varepsilon_{mt})$. And let $G(n|X_{mt}, \varepsilon_{mt})$ be the distribution of the number of active firms in equilibrium conditional on $(X_{mt}, \varepsilon_{mt})$ and given that one of the firms is active with probability one. (i) Obtain the expression of the probability distribution $G(n|X_{mt}, \varepsilon_{mt})$ in terms of the probability of entry $P(X_{mt}, \varepsilon_{mt})$. (ii) Derive the expression for the expected profit of an active firm in terms of the probability of entry. (iii) Obtain the expression of the equilibrium mapping that defines implicitly the equilibrium probability of entry $P(X_{mt}, \varepsilon_{mt})$.

NOTE: *For Questions 6 and 7, consider the following approximation to the function* $\ln E(v(n_{mt}) \mid X_{mt}, \varepsilon_{mt}, 1$ *sure):*

$$\ln E(v(n_{mt})|X_{mt}, \varepsilon_{mt}, 1\text{sure}) \simeq \ln v(1) + \sum_{n=1}^{N_{mt}} G(n|X_{mt}, \varepsilon_{mt}) \; \left[\frac{v(n) - v(1)}{v(1)} \right]$$

This is a first order Taylor approximation to $\ln E(v(n_{mt})|X_{mt}, \varepsilon_{mt}, 1sure)$ *around the values* $v(1) = v(2) = \ldots = v(N)$*, i.e., no competition effects. The main advantage of using this approximation for estimation is that it is linear in the parameters* $\left[\frac{v(n)-v(1)}{v(1)}\right]$.

Question 6. [20 points] Suppose that $\varepsilon_{mt} \equiv \varepsilon^S_{mt} - \varepsilon^F_{mt}$ is just an aggregate time effect, $\varepsilon_{mt} = \gamma_t$. Use a two-step pseudo maximum likelihood method to estimate the vector of parameters:

$$\theta \equiv \left\{ \frac{1}{\sigma_\xi}, \; \frac{\beta^S_1}{\sigma_\xi}, \; \frac{\ln v(1) + \beta^S_0 - \beta^F}{\sigma_\xi}, \; \frac{v(n) - v(1)}{\sigma_\xi \; v(1)} : n = 2, 3, \ldots \right\}$$

for each of the five industries separately. Given these estimates, obtain estimates of the parameters $\frac{v(n+1)}{v(n)}$ for $n \in \{1, 2, \ldots, n_{\max}\}$. Present a figure of the estimated function $\frac{v(n+1)}{v(n)}$ for each of the five industries. Interpret the results. Based on these results, what can we say about *the nature of competition* in each of these industries? Compare these results to those from the estimation of the *BR-91* models in Questions 2 and 3.

Question 7. [10 points] Repeat the same exercise as in Question 7 but using the following specification of the unobservable ε_{mt}:

$$\varepsilon_{mt} = \gamma_t + \delta_m$$

where γ_t are time effects that can be captured by using time-dummies; and δ_m are fixed market effects that can be captured by using market-dummies. Comment the results. Compare these results to those in Questions 2, 3, and 6.

Final exam 2010

INDUSTRIAL ORGANIZATION II (ECO 2901)

University of Toronto. Department of Economics. Spring 2010
Instructor: Victor Aguirregabiria

FINAL EXAM

Take Home Exam. Due on Tuesday, April 13 before Midnight

Answer all the questions.

Consider an oligopoly industry characterized by local competition. A researcher has panel data of M local markets over T years, where M is large and T is small. Markets are indexed by m and years are indexed by t. For every market and year, the dataset includes information on: market size, s_{mt}; the number of active firms, n_{mt}; the number of new entrants during the year, en_{mt}; and the number of exiting firms, ex_{mt}. A descriptive analysis of these data reveals the following *stylized facts*.

(SF.1) There is simultaneous entry and exit at the individual market level. A significant proportion of the observations (m,t) are characterized by $en_{mt} > 0$ and $ex_{mt} > 0$.
(SF.2) For every cross-section of markets (every period t), there is positive correlation between the number of entrants and the number of exiting firms.
(SF.3) Conditional on market size s_{mt}, entry is positively correlated (and exit is negatively correlated) with the number of incumbent firms at the beginning of the year. For instance, in a linear regression of en_{mt} on s_{mt} and n_{mt-1} the estimate of the coefficient associated to n_{mt-1} is positive and statistically significant.

Consider the following model of oligopoly competition in a local market. There are N firms that may operate in the market. A firm in this market can be either active or inactive. The profit of an inactive firm is zero. The profit of an active firm in a market with n competitors is:

$$\Pi_{mt}(n) = s_{mt}\left(\theta_0^{VP} - \theta_1^{VP} n\right) - \theta^{FC} - \varepsilon_{imt} - (1 - a_{imt-1})\theta^{EC}$$

θ_0^{VP}, θ_1^{VP}, θ^{FC}, and θ^{EC} are parameters. a_{imt-1} is the binary indicator of the event "firm i was an incumbent at period $t-1$". ε_{imt} is a component of the fixed operating cost that varies over time, across markets, and across firms, and it is private information of firm i. We assume that ε_{imt} is iid over time, markets, and firms, with a $N(0, \sigma_\varepsilon^2)$ distribution. Market size evolves exogenously over time according to a Markov process with transition probability function $f_s(s_{mt+1}|s_{mt})$. Every period t, firms observe market size, the number of active firms in the market at previous period, and their own private fixed cost, and then they decide simultaneously whether to be active in the market or not. Firms are forward-looking and play strategies that depend only on payoff-relevant state variables. The equilibrium in this model is a Markov Perfect Equilibrium (MPE). Given that firms are identical, up to their private information ε_{imt}, we consider only symmetric MPE.

Question 1 (20 points): Describe in detail the structure of a MPE in this model. Derive and explain the following objects in this model:

(1.1) the vector of payoff relevant state variables;
(1.2) the expected one-period profit;
(1.3) the transition probability of the state variables;
(1.4) the dynamic decision problem of an incumbent firm and his best response function;
(1.5) the dynamic decision problem of a potential entrant and his best response function;
(1.6) the best response probability function;
(1.7) the MPE as a fixed point of a mapping in the space of firm's choice probabilities.

Question 2 (10 points): Let x_{mt} be the vector (s_{mt}, n_{mt-1}). Let $P_0(x_{mt})$ be the probability that a potential entrant chooses to enter in the market, and let $P_1(x_{mt})$ be the probability that an incumbent firm decides to stay in the market. Let $P_{en}(en_{mt}|x_{mt})$ and $P_{ex}(en_{mt}|x_{mt})$ be the probability distributions for the number of entrants and the number of exits conditional on x_{mt}, respectively.

(2.1) Write the distribution $P_{en}(.|x_{mt})$ in terms of the probability $P_0(x_{mt})$, and the distribution $P_{ex}(.|x_{mt})$ in terms of the probability $P_1(x_{mt})$;
(2.2) Show that there is a one-to-one relationship between $P_{en}(.|x_{mt})$ and $P_0(x_{mt})$, and between $P_{ex}(.|x_{mt})$ and $P_1(x_{mt})$.
(2.3) Based on the result in (2.2), define a MPE in the model as a fixed point of a mapping in the space of the probability distributions $P_{en}(.|x_{mt})$ and $P_{ex}(.|x_{mt})$.

Question 3 (20 points): Consider the conditional log-likelihood function:

$$l(\theta) = \sum_{m=1}^{M} \log \Pr \left(n_{m2},\, n_{m3},\, ...,\, n_{mT} \mid n_{m1},\, s_{m1},\, s_{m2},\, ...,\, s_{mT} \right)$$

where θ is the vector of structural parameters.

(3.1) Write this log-likelihood function in terms of the probabilities $P_{en}(en_{mt}|x_{mt})$ and $P_{ex}(ex_{mt}|x_{mt})$.
(3.2) Suppose that for every value of θ the model has a unique equilibrium. Describe in detail a method for the estimation of θ in this model.
(3.3) In general, there are values of θ for which the model has multiple equilibria. Describe in detail a two-step method for the estimation of θ. Explain how this method can be extended recursively.

Question 4 (10 points): Explain why this model can explain the empirical evidence in (SF.1) but it cannot explain stylized facts (SF.2) and (SF.3).

Question 5 (20 points): To explain the evidence in (SF.2) consider the following two hypotheses. Hypothesis 1 (Market heterogeneity in the variance of idiosyncratic shocks): Markets are heterogeneous in the dispersion of the private information shocks. For instance, $\varepsilon_{imt} \sim N(0, \sigma^2_{mt})$ where $\sigma^2_{mt} = (\lambda s_{mt})^2$. Hypothesis 2 (Creative Destruction): a firm's idiosyncratic shock ε_{imt} has two components, $\varepsilon_{imt} = \varepsilon^{(p)}_{imt} + \varepsilon^{(c)}_{imt}$, where $\varepsilon^{(p)}_{imt}$ is private information of firm i but $\varepsilon^{(c)}_{imt}$ is common knowledge of all the firms.

(5.1) Explain why these hypotheses could explain the evidence in (SF.2).

(5.2) Is it possible to distinguish empirically between the two hypotheses using these data? Explain why/how.

(5.3) Propose a method to estimate the model under hypothesis 1.

(5.4) Propose a method to estimate the model under hypothesis 2.

Question 6 (20 points): To explain the evidence in (SF.3) consider the following two hypotheses. Hypothesis 3 (Market heterogeneity in average fixed costs): The fixed operating cost in market m is $FC_m = \theta^{FC} + \omega_m$, where ω_m is a zero mean random variable that is common knowledge to all the firms. Hypothesis 4 (Uncertainty with "learning-by-being-active" and "learning from others"): The fixed operating cost in market m is $FC_m = \theta^{FC} + \xi_m$, and ξ_m is a zero mean random variable that is unknown to a potential entrant but it is perfectly known by active firms. If a firm enters in market m, it immediately learns the value of ξ_m, i.e., learning-by-being-active. Potential entrants observe whether incumbent firms stay in the market or exit, and they use this information to update their beliefs about the value of ξ_m, i.e., learning from others.

(6.1) Explain why these hypotheses could explain the evidence in (SF.3).

(6.2) Is it possible to distinguish empirically between the two hypotheses using these data? Explain why/how.

(6.3) Propose a method to estimate the model under hypothesis 3. For simplicity, suppose that ω_m can take only two values and it has a known distribution.

(6.4) Propose a method to estimate the model under hypothesis 4. For simplicity, suppose that ξ_m can take only two values and it has a known distribution.

Problem set 2011

Industrial Organization II (ECO 2901)
Winter 2011. Victor Aguirregabiria

Problem Set #1
Model of Industry Dynamics with Homogeneous Firms
Due on Thursday, March 3, 2011

Context. At the end of year 2002, the federal government of Greenchistan introduced a new environmental regulation on the cement industry, one of the major polluting industries. The most important features of this regulation is that new plants, in order to operate in the industry, should pass an environmental test and should install a piece of equipment that contributes to reduce pollutant emissions. Industry experts consider that this new law increased the sunk cost of entry in the industry. However, these experts disagree in the magnitude of the increase in sunk costs. There is also disagreement with respect to whether the new law affected production costs, competition, prices, and output. You have been hired by the Ministry of Industry as an independent researcher to study and to evaluate the short-run and long-run effects of this policy on output, prices, firms' profits, and consumer welfare.

Data. To perform your evaluation, you have a panel dataset with annual information on the industry for the period 1998-2007. The Stata datafile eco2901_problemset_01_dynamicbr_2011.dta contains panel data from 1000 local markets (census tracts) over 10 years (1998-2007) for the cement industry of Greenchistan. The local markets in this dataset have been selected following criteria similar to the ones in Bresnahan and Reiss (1991). This is the list of variables in the dataset:

Variable name		Description
market	:	Code of local market
year	:	Year
pop	:	Population of local market (in # people)
pcincome	:	Per capita income in local market
output	:	Annual output (tons of cement) produced in the local market
price	:	Price of cement in local market
pinput	:	Price index of intermediate inputs in local market
nplant	:	Number of cement plants in local market at current year
nplant_1	:	Number of cement plants in local market at previous year

Model. To answer our empirical questions we consider a model similar to the dynamic version of Bresnahan and Reiss model that we have seen in class (Bresnahan and Reiss, AES

1994). The main difference with respect to that model is that we specify the demand function and the cost function in the industry and make it explicit the relationship between these primitives and the value of a plant.

Demand of cement in market m at period t. We assume that cement is an homogeneous product and consider the following inverse demand function:

$$\ln P_{mt} = \alpha_0^D + \alpha_1^D \ln POP_{mt} + \alpha_2^D \ln PCINC_{mt} - \alpha_3^D \ \ln Q_{mt} + \varepsilon_{mt}^D$$

where $\alpha^{D\prime}s$ are demand parameters, Q_{mt} represents output, POP_{mt} is population, $PCINC_{mt}$ is per capita income, P_{mt} is price, and ε_{mt}^D is a component of the demand that is unobserved to the researcher.

Production costs. Let q_{mt} be the amount of output of a cement plant in market m and period t. And let X_{mt} be the vector of observable state variables $X_{mt} = (n_{mt-1},\ POP_{mt},\ PCINC_{mt},\ PINPUT_{mt})$, where $PINPUT_{mt}$ is the index price of inputs (energy and limestone). The production cost function is $C_{mt}(q_{mt}) = FC_{mt} + VC_{mt}(q_{mt})$, where FC_{mt} and $VC_{mt}(q_{mt})$ are the fixed cost function and the variable cost function, respectively:

$$\begin{aligned} FC_{mt} &= X_{mt}\ \alpha^{FC} + \varepsilon_{mt}^{FC} \\ VC_{mt}(q_{mt}) &= (X_{mt}\ \alpha^{MC} + \varepsilon_{mt}^{MC})\ q_{mt} + \frac{\delta}{2}(q_{mt})^2 \end{aligned}$$

where α^{FC} and α^{MC} are vectors of parameters, δ is a parameter that captures economies/diseconomies to scale, and ε_{mt}^{FC} and ε_{mt}^{MC} are components of the fixed cost and the marginal cost, respectively, that are unobserved to the researcher. There are not good reasons to believe that production costs depend on the number of firms at previous period, n_{mt-1}, but in principle we can allow for that dependence.

Entry costs and scrapping value. Let EC_{mt} be the sunk entry cost in market m at period t. We normalize the scrapping value of a plant to zero. The sunk cost may depend on market characteristics and on the number of firms.

$$EC_{mt} = X_{mt}\ \alpha^{EC} + \varepsilon_{mt}^{EC}$$

where α^{EC} is a vector of parameters and ε_{mt}^{EC} is unobserved for the researcher.

Unobservables. Let ε_{mt} be the vector of unobservables $\varepsilon_{mt} \equiv (\varepsilon_{mt}^D,\ \varepsilon_{mt}^{MC},\ \varepsilon_{mt}^{FC},\ \varepsilon_{mt}^{EC})$. We allow for serial correlation in these unobservables. In particular, we assume that each of these unobservables follows an AR(1) process. For $j \in \{D,\ MC,\ FC,\ EC\}$:

$$\varepsilon_{mt}^j = \rho^j\ \varepsilon_{mt-1}^j + u_{mt}^j$$

where $\rho^j \in [0,1)$ is the autorregressive parameter, and the vector $u_{mt} = (u_{mt}^D, u_{mt}^{MC}, u_{mt}^{FC}, u_{mt}^{EC})$ is i.i.d. over markets and over time with a joint normal distribution with zero means and variance-covariance matrix Ω.

Question 1 [15 points]. (a) Propose an estimator of the demand parameters and the explain the assumptions under which the estimator is consistent. (b) Obtain estimates and standard errors. (c) Test the null hypothesis of "no structural break" in demand parameters after year 2002.

Question 2 [15 points]. (a) Describe how to use the Cournot equilibrium conditions to estimate the parameters in the variable cost function. Explain the assumptions under which the estimator is consistent. (b) Obtain estimates and standard errors. (c) Test the null hypothesis of "no structural break" in the variable cost parameters after year 2003.

Question 3 [20 points]. Assume that $\rho^{FC} = \rho^{EC} = 0$. (a) Describe how to estimate the parameters in the entry cost and in the fixed cost functions. Show that these costs are identified in dollar amounts (i.e., not only up to scale). Explain the assumptions under which the estimator is consistent. (b) Obtain estimates and standard errors. (c) Test the null hypothesis of "no structural break" in the entry cost and fixed cost parameters after year 2003.

Now, we use our estimates to evaluate the effects of the policy change. First, we want to obtain the steady-state distribution of the number of firms before and after the policy change. For simplicity, we consider this evaluation for a "median market" and assume that most of the exogenous state variables are constant over time. Consider a local market with the median values of the exogenous variables in the vector X_{mt} and of the shocks ε_{mt}^D, ε_{mt}^{MC}, and ε_{mt}^{EC}, and suppose that these variables stay constant forever at these median values such that the only exogenous variable that changes over time is ε_{mt}^{FC}. Let $V^{Bef}(n_{t-1}) - \varepsilon_t^{FC}$ and $V^{Aft}(n_{t-1}) - \varepsilon_t^{FC}$ be the value functions of an incumbent firm in this "median market" before and after the policy change, respectively.

Question 4 [30 points]. (a) Compute the functions $V^{Bef}(n_{t-1})$ and $V^{Aft}(n_{t-1})$; (b) Given the functions $V^{Bef}(n_{t-1})$ and $V^{Aft}(n_{t-1})$, obtain the transition probabilities for the number of firms "Before" and "After" the policy change: $P^{Bef}(n_t|n_{t-1})$ and $P^{Aft}(n_t|n_{t-1})$; (c) Given these transition probabilities, obtain the steady-state distribution of the number of firms "Before" and "After" the policy change: $p_*^{Bef}(n)$ and $p_*^{Aft}(n)$.

[Hint: To obtain this steady-state distributions, we can use a simple iterative procedure. Let $\{0, 1, ..., N\}$ be the set of possible values of n_t, where N is a large value. Let P be the transition probability matrix such that the n-th column of this matrix contains the column

vector of probabilities $(P(0|n),\ P(1|n),\ ...,\ P(N|n))'$. And pet p_* be the column vector with the steady-state distribution $(p_*(0),\ p_*(1),\ ...,\ p_*(N))'$. Then, by definition $p_* = Pp_*$. This defines p_* as the solution of this fixed-point mapping under the constraint $\sum_{n=0}^{N} p_*(n) = 1$. A straightforward way of computing p_* is by successive iterations in the fixed point mapping, starting with an initial p_* that satisfies the restriction $\sum_{n=0}^{N} p_*(n) = 1$. It is simple to show that this mapping is a contraction, so it has a unique fixed point and it always converges, and also by construction the constraint $\sum_{n=0}^{N} p_*(n) = 1$ is always satisfied if P is a well-defined transition matrix].

Question 5 [20 points]. Given the steady state distribution from Question 4, obtain the "Before" and "After" steady-state distributions of: (a) aggregate output; (b) price; (c) firms' profits; and (d) consumer welfare. Comment the results. According to these results, which are the most important effects of this policy.

Final exam 2011

INDUSTRIAL ORGANIZATION II (ECO 2901)

University of Toronto. Department of Economics. Spring 2011
Instructor: Victor Aguirregabiria

FINAL EXAM

Monday, April 18, 2011. From 9:00-12:00 (3 hours)

INSTRUCTIONS: The exam consists of 5 Questions (with sub-questions). You have to answer all the questions. No study aids, including calculators, are allowed.
TOTAL MARKS = 100

Consider the retail industry of coffee shops in a region. This industry is characterized by the leadership of three retail chains that we denote as SB, SC, and TH. You may think in *Starbucks*, *Second Cup* and *Tim Hortons*, though this problem deals with an hypothetical industry. Suppose that the retail chains SC and TH have announced a merger. You have been hired by the *Competition Commission* to evaluate the effects of this merger (in the hypothetical case that it is approved) on prices, market shares, profits, and consumer welfare.

You have been provided with a panel dataset with information from this industry that covers $T = 30$ quarters and $M = 500$ local markets (census blocks). We index time by t, markets by m, and firms by i. The information in the dataset includes: prices, p_{imt}; quantities, q_{imt}; a measure of market size, h_{mt}; average household income, y_{mt}; rental prices, r_{mt}; and average wage in the retail sector, w_{mt}. Of course, the dataset includes only 'pre-merger' information.

To evaluate the effects of the merger, you propose and estimate a structural model of competition in this industry. Firms compete in local markets, and competition is independent across local markets. The model of competition in a single market has the following features. Every quarter t, firms decide simultaneously whether to have or not a store in the market. This decision is static (i.e., there are not sunk costs of entry). Then, the active firms in the local market compete in prices ala Nash-Bertrand, and this competition determines firms' profits. The profit of firm i in market m is:

$$\Pi_{imt} = a_{imt} \ \left[(p_{imt} - MC_{imt}) \ q_{imt} - FC_{imt} \right]$$

$a_{imt} \equiv 1\{q_{imt} > 0\}$ is the binary indicator of the event "firm i has a store in market m at quarter t". And MC_{imt} and FC_{imt} are the marginal cost and the fixed cost of firm i in market m, respectively. Firms' products are differentiated. We model consumer demand using a logit model where product 'quality' can interact with consumer income at the market level. The market share of firm i in market m is:

$$s_{imt} \equiv \frac{q_{imt}}{h_{mt}} = \frac{a_{imt} \ \exp\{\delta_{imt}\}}{1 + \sum_j a_{jmt} \ \exp\{\delta_{jmt}\}}$$

with

$$\delta_{imt} = \alpha_i^{(1)} + \alpha_i^{(2)} y_{mt} - \alpha_i^{(3)} p_{imt} - \alpha_i^{(4)} y_{mt}\ p_{imt} + \xi_m^{(1)} + \xi_t^{(2)} + \xi_{imt}^{(3)}$$

where $\{\alpha_i^{(1)},\, \alpha_i^{(2)},\, \alpha_i^{(3)},\, \alpha_i^{(4)} : i = SB,\, SC,\, TH\}$ are demand parameters, and $\xi' s$ represent error terms that are observable to firms but unobservable to you as a researcher. The specification of marginal costs is:

$$MC_{imt} = \beta_i^{(1)} + \beta_i^{(2)} w_{mt} + v_m^{(1)} + v_t^{(2)} + v_{imt}^{(3)}$$

where $\{\beta_i^{(1)},\, \beta_i^{(2)} : i = SB,\, SC,\, TH\}$ are parameters, and $v' s$ represent error terms that are observable to firms but unobservable to you as a researcher. Finally, the specification of fixed operating costs is:

$$FC_{imt} = \gamma_i^{(1)} + \gamma_i^{(2)} r_{mt} + \varepsilon_m^{(1)} + \varepsilon_t^{(2)} + \varepsilon_{imt}^{(3)}$$

where $\{\gamma_i^{(1)},\, \gamma_i^{(2)} : i = SB,\, SC,\, TH\}$ are parameters, and $\varepsilon' s$ represent error terms that are observable to firms but unobservable to you as a researcher.

As for the unobservable variables of the structural model, we make the following assumptions. The variables $\xi_m^{(1)}$, $v_m^{(1)}$, and $\varepsilon_m^{(1)}$ are treated as market fixed effects and controlled for by including market dummies. The variables $\xi_t^{(2)}$, $v_t^{(2)}$, and $\varepsilon_t^{(2)}$ are treated as time 'fixed effects' and controlled for by including time dummies. And the variables $\xi_{imt}^{(3)}$, $v_{imt}^{(3)}$, and $\varepsilon_{imt}^{(3)}$ are assumed independently distributed of (exogenous) observed market characteristics, h_{mt}, y_{mt}, w_{mt}, and r_{mt}.

−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−

Question 1.1. (20 points). Estimation of Demand. The demand model can be described by the equations:

$$\ln(s_{imt}/s_{0mt}) = \delta_{imt} \qquad \text{if } \ a_{imt} = 1$$

where s_{0mt} is the share of the outside good, $s_{0mt} = 1 - s_{SBmt} - s_{SCmt} - s_{THmt}$.

(a) Discuss the endogeneity problems (both endogenous prices and endogenous entry) in the estimation of demand parameters in this model.
(b) Propose a method for the estimation of the demand parameters that deals with these endogeneity problems. Explain your method in detail.
(c) Suppose that we assume that the error terms $\xi_{imt}^{(3)}$ are unknown to firms when they decide to be active or not in the market. Explain how this assumption simplifies the estimation of demand parameters.

Question 1.2. (20 points). Estimation of Marginal Costs. Suppose that you have consistent estimates of demand parameters, including market and time fixed effects. Nash-Betrand competition implies the following best response functions for prices:

$$p_{imt} = MC_{imt} + \frac{1}{(\alpha_i^{(3)} + \alpha_i^{(4)} y_{mt})\ (1 - s_{imt})} \qquad \text{if } \ a_{imt} = 1$$

(a) In the estimtation of marginal cost parameters, discuss the selection or endogeneity problem due to endogenous firm entry.

(b) Propose a method for the estimation of marginal cost parameters that deals with this endogeneity problem. Explain your method in detail.

(c) Suppose that we assume that the error terms $v^{(3)}_{imt}$ are unknown to firms when they decide to be active or not in the market. Explain how this assumption simplifies the estimation of demand parameters.

Question 1.3. (20 points). Construction of Variable Profits. Suppose that you have consistent estimates of demand and marginal cost parameters. Let $VP_{imt}(a_{SB}, a_{SC}, a_{TH})$ be the variable of firm i in market m at period t under the hypothetical market structure (a_{SB}, a_{SC}, a_{TH}).

(a) Explain in detail how to calculate estimated values of variable profit function $VP_{imt}(a_{SB}, a_{SC}, a_{TH})$ for every market-quarter in the sample and for every hypothetical market structure $(a_{SB}, a_{SC}, a_{TH}) \in \{0,1\} \times \{0,1\} \times \{0,1\}$. Assume that $\xi^{(3)}_{imt} = v^{(3)}_{imt} = 0$ for every (i, m, t).

(b) Explain why the assumption $\xi^{(3)}_{imt} = v^{(3)}_{imt} = 0$ for every (i, m, t) helps in the calculation of $VP_{imt}(a_{SB}, a_{SC}, a_{TH})$.

(c) Suppose that $\xi^{(3)}_{imt}$ and $v^{(3)}_{imt}$ are not zero but we still assume that they are unknown to firms when they make their entry decision. Suppose that $\xi^{(3)}_{imt}$ and $v^{(3)}_{imt}$ are iid over (i, m, t). Now, $VP_{imt}(a_{SB}, a_{SC}, a_{TH})$ represents expected variable profit, where the expectation is taken over the distribution of $\xi^{(3)}_{imt}$ and $v^{(3)}_{imt}$. Explain how to estimate this expected variable profit.

Question 1.4. (20 points). Estimation of Fixed Costs. Suppose that you have consistent estimates of the variable profit function $VP_{imt}(a_{SB}, a_{SC}, a_{TH})$ for every firm i, market, and time period, and for every possible market structure (a_{SB}, a_{SC}, a_{TH}). The next step is the estimation of parameters in fixed costs. Suppose that the variables $\varepsilon^{(3)}_{imt}$ are firms' private information shocks that are independent across firms and over time and $\varepsilon^{(3)}_{imt}$ is iid extreme value type 1 with dispersion parameter σ_i. Given beliefs about the entry strategies of the other firms, firm i's best response is:

$$a_{imt} = 1\{\, E\left(VP_{imt}(1, a_{-imt}) \mid x_{mt}\right) - FC_{imt} \geq 0\}$$

x_{mt} is the vector of exogenous market characteristics of market m at period t, including h_{mt}, y_{mt}, w_{mt}, r_{mt}, and the estimated fixed effects $\xi^{(1)}_m$, $\xi^{(2)}_t$, $v^{(1)}_m$, and $v^{(2)}_t$. $E\left(VP_{imt}(1, a_{-imt}) \mid x_{mt}\right)$ is the expected variable profit of firm i if the firm is active in the market and integrated over the unknown private information of the other firms.

(a) Let $P_i(x_{mt})$ be the Conditional Choice Probability (CCP) that represents $\Pr(a_{imt} = 1|x_{mt})$. Show how to represent a Bayesian Nash Equilibrium (BNE) of the entry game as system of 3 equations with the 3 unknowns $P_{SB}(x_{mt})$, $P_{SC}(x_{mt})$, and $P_{TH}(x_{mt})$. Write the functional form of this system of equations.

(b) Explain how to compute a BNE in a market m at period t.

(c) Explain in detail a method to estimate the fixed cost parameters in this model of entry.

(d) Are the parameters $\gamma_i^{(1)}$, $\gamma_i^{(2)}$, and σ_i separately identified? Why/Why not?

(e) Explain why the assumption that $\varepsilon_{imt}^{(3)}$ are independent private information shocks facilitate the identification and estimation of the model.

Question 1.5. (20 points). Counterfactual experiment: Merger. Suppose that you have consistent estimates of all the parameters of the model.

(a) Explain how to compute firms' profits and consumer surplus for every market-quarter observation in the data.

Suppose that retail chains SC and TH merge to become a single corporation but with two different brands: brand SC and brand TH. Suppose that the brand-specific parameters in demand and costs remain the same after the merger. The only differences between pre-merger and post-merger competition are: (1) the new firm chooses prices of SC and TH to maximize the total variable profits of the company; and (2) the new firm chooses market entry decisions, a_{SC} and a_{TH}, to maximize the total profits of the company.

(b) Explain in detail the different steps to calculate firms' profits and consumer surplus under this counterfactual post-merger scenario for every market-quarter observation in the data.

Problem set 2012

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:

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

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.

Final exam 2012

INDUSTRIAL ORGANIZATION II (ECO 2901)

University of Toronto. Department of Economics. Winter 2012
Instructor: Victor Aguirregabiria

FINAL EXAM: April 16, 2012. From 9:00-12:00 (3 hours)

INSTRUCTIONS: The exam consists of three Problems. You have to answer all the questions. No study aids, including calculators, are allowed.

TOTAL MARKS = 100

PROBLEM 1 (40 points). Answer the following questions on the article *"Automobile Prices in Market Equilibrium,"* by Berry, Levinshon, and Pakes (*Econometrica*, 1995).

Question 1.1 (10 points). Write the regression equation that relates market shares with the average indirect utility of product j. Explain how to obtain this equation.

Question 1.2 (10 points). Under the assumption of Bertrand competition, obtain the equation that relates the equilibrium price-cost-margin of product j with demand and demand elasticities. For simplicity, assume that each firm produces a single product.

Question 1.3 (10 points). Describe the instrumental variables approach proposed by BLP to estimate demand and supply parameters. More specifically, explain: (a) moment conditions; (b) assumptions for the validity of instruments; (c) sample criterion function minimized by the estimator. You do not have to describe here the algorithm for the computation of the estimator.

Question 1.4 (10 points). Explain the main challenges in the computation of the estimator in Question 1.3. Describe the Nested Fixed Point algorithm in this model.

PROBLEM 2 (30 points). Answer the following questions on the article *"Measuring the Implications of Sales and Consumer Inventory Behavior,"* by Hendel and Nevo (*Econometrica*, 2006).

Question 2.1 (10 points). Explain why a static model that ignores dynamics in the demand of a storable good can provide biased estimates of long-run price elasticities of demand.

Question 2.2 (10 points). Suppose a simplified version of the model in Hendel and Nevo where the storable good is not differentiated (homogeneous product). Also, suppose that the dataset used by Hendel and Nevo included data on households' inventories of the product at the end of each month, e.g., every household participating in the survey should maintain a record of its inventory of laundry detergent at the last day of each month. Propose an approach to estimate the long-run price elasticity of demand using this model and these data.

Question 2.3 (10 points). Describe the complexities that product differentiation and unobserved household inventories incorporate in the estimation of long-run demand elasticities.

PROBLEM 3 (30 points). Answer the following questions on the article *"The Costs of Environmental Regulation in a Concentrated Industry,"* by Ryan (*Econometrica*, 2012).

Question 3.1 (10 points). Propose a Difference-in-Differences regression approach to evaluate the policy question in this paper. Describe the dependent and the explanatory variables of the equation(s), and the control and experimental groups. Discuss the relative merits and limitation of this approach relative to Ryan's approach.

Question 3.2 (20 points). Consider Ryan's approach to the estimation of the parameters in variable the variable cost function.

(a) (5 points). Describe his assumption(s) about unobserved firm heterogeneity in marginal costs.
(b) (5 points). Suppose that this maintained assumption(s) is not true. Which are the potential implications on the estimation of variable profits and on the policy evaluation?
(c) (5 points). Propose a method to test this assumption.
(d) (5 points). Propose a method that relaxes this assumption.

Problem set 2013

Industrial Organization II (ECO 2901)

Winter 2013. Victor Aguirregabiria

Problem Set #1

Due of Friday, March 22, 2013

TOTAL NUMBER OF POINTS: 200

PROBLEM 1 [30 points]. Consider the estimation of a model of demand of differentiated products using aggregate market data as in Berry (1994) and Berry, Levinsohn, and Pakes (1995). The dataset includes information on prices (p), quantities (q), and product characteristics other than price (X) for J products over M separate markets:

$$\text{Data} = \{p_{jm},\, q_{jm},\, X_{jm} : j = 1, 2, ..., J;\, m = 1, 2, ..., M\}$$

where we index markets by m and products by j. In this dataset the number of markets M is large relative to the number of product varieties, e.g., $J = 50$ and $M = 1,000$. The specification of the model is the one in the random-coefficients 'BLP' model, where the utility of buying product j for consumer i in market m is:

$$V_{ijm} = X_{jm} \ \left[\beta + v_{im}^{\beta}\right] - p_{jm} \ \left[\alpha + v_{im}^{\alpha}\right] + \xi_{jm} + \varepsilon_{ijm}$$

where v_{im}^{β} and v_{im}^{α} are zero mean normal random variables that capture consumer heterogeneity in the marginal utility of product characteristics, and ε_{ijm} is a type 1 extreme value distributed variable that also captures consumer heterogeneity in preferences. For the outside alternative, $j = 0$, we have that $V_{i0m} = 0$. H_m is the number of consumers in market m, i.e., market size.

Question 1.1 [10 points] Suppose that any observable measure of market size H_m available to the researcher includes substantial measurement error. Propose a simple approach to deal with this problem. Explain in detail your proposed method.

Question 1.2 [10 points] Suppose that a substantial proportion of products are not available in all the M markets. For instance, the top-5 products (according to their market shares at the national level) are available in 95% of the local markets, while products below the top-20 are available only in 60% of the local markets. There are multiple factors that contribute to explain why a product is available or not in a local market, e.g., market size, competition, local consumer preferences, distance to production size, economies of density, etc. We believe that in the industry under study an important factor to explain these differences in product availability across markets has to do with heterogeneity among local markets in the preferences of the average local consumer, as represented by the unobserved variables $\{\xi_{jm}\}$.

(a) Discuss the implications of this issue on the properties of the standard GMM estimator using BLP moment conditions.

(b) Propose an approach to deal with this problem. Explain in detail your proposed method.

Question 1.3 [10 points] Recently, Petrin and Train (Journal of Marketing Research, 2010) and Kim and Petrin (WP, 2011) have proposed Control Function (CF) approaches to estimate the 'BLP' model and extensions of this model that allow for interactions between market level unobservables ξ and price p in the utility function. This CF is in the spirit of Rivers and Vuong (JE, 1988) at it operates in two-steps. The first step is an OLS estimation of a linear regression for the reduced form equation of prices. In the second step, the residuals from the first-step regression are plugged-in utility function to control the unobservables $\{\xi_{jm}\}$ and the the parameters of the model are estimated by Maximum Likelihood.

(a) Discuss in more detail the CF approach providing specific equations and formulas.

(b) Discuss the relative advantages and limitations of the CF approach versus the GMM-BLP approach.

PROBLEM 2 [10 points]. Describe in detail Ackerberg-Frazer-Caves (2006) criticism to the identification of the parameters in the Cobb-Douglas Production Function using Olley-Pakes Control Function approach.

PROBLEM 3 [30 points]. Consider the Two-Players Binary Choice Probit Game of complete information in Tamer (REStud, 2003). The structural equations of the model are the following best response functions:

$$\begin{aligned} Y_1 &= 1\{\ \alpha_1' X + \beta_1\ Z_1 - \delta_1\ Y_2 - \varepsilon_1 \geq 0\ \} \\ Y_2 &= 1\{\ \alpha_2' X + \beta_2\ Z_2 - \delta_2\ Y_1 - \varepsilon_2 \geq 0\ \} \end{aligned}$$

where: $Y_1 \in \{0,1\}$ and $Y_2 \in \{0,1\}$ represent players' decisions; α_1, α_2, β_1, β_2, δ_1, and δ_2 are parameters, and we assume that $\delta_1 \geq 0$ and $\delta_2 \geq 0$; X, Z_1, and Z_2 are exogenous observable variables; and ε_1 and ε_2 are Normal random variables independent of (X, Z_1, Z_2) with zero mean, unit variances, and correlation parameter ρ. We use the $\Phi^{(2)}(\varepsilon_1, \varepsilon_2; \rho)$ to represent the CDF of $(\varepsilon_1, \varepsilon_2)$. The researcher observes a random sample of M markets with information on $\{Y_{1m},\ Y_{2m},\ X_m,\ Z_{1m},\ Z_{2m} : m = 1, 2, ..., M\}$. We are interested in using this sample to estimate the vector of structural parameters $\theta = (\alpha_1,\ \alpha_2,\ \beta_1,\ \beta_2,\ \delta_1,\ \delta_2,\ \rho)'$. For $(y_1, y_2) \in \{(0,0),\ (0,1),\ (1,0),\ (1,1)\}$, define the *Conditional Choice Probability* (CCP) function

$$P(y_1, y_2 \mid x, z_1, z_2; \theta) = \Pr\left(Y_1 = y_1,\ Y_2 = y_2 \mid X = z,\ Z_1 = z_1,\ Z_2 = z_2,\ \theta\right)$$

Question 3.1 [10 points]. Obtain the reduced form equations of the model, i.e., the relationship between the four possible values of the endogenous variables (Y_1,Y_2) and the exogenous variables and parameters.

Question 3.2 [5 points]. Using the reduced form equations, obtain the expressions for the CCPs $P(0,0 \mid x, z_1, z_2; \theta)$ and $P(1,1 \mid x, z_1, z_2; \theta)$. Use $\Phi^{(2)}(\varepsilon_1, \varepsilon_2; \rho)$ to represent the join CDF of $(\varepsilon_1, \varepsilon_2)$.

Question 3.3 [15 points]. Suppose that: (i) $\beta_1 \neq 0$ and $\beta_2 \neq 0$; and (ii) the distribution of (Z_1,Z_2) is such that the support set is $\mathbb{R}^2$, in the limit as $Z_1 \rightarrow \infty$ we have that the distribution of (X, Z_2) is non-degenerate, and similarly in the limit as $Z_2 \rightarrow \infty$ the distribution of (X, Z_1) is non-degenerate. For instance, condition (ii) is satisfied if (Z_1,Z_2) are jointly normally distributed conditional on Z. Prove formally that under conditions (i) and (ii) θ is identified using the data and CCP functions $P(0,0 \mid x, z_1, z_2; \theta)$ and $P(1,1 \mid x, z_1, z_2; \theta)$. [Hint: Read the proof in the Appendix of Tamer (2003)).

PROBLEM 4 [30 points]. Consider a Two-Player Game of Market Entry with Incomplete Information. The players' payoff functions are:

$$\begin{aligned}
\Pi_1 &= \pi_1^M(X, Z_1) + Y_2 \, \left[\pi_1^D(X, Z_1) - \pi_1^M(X, Z_1)\right] - \varepsilon_1 \\
\Pi_2 &= \pi_2^M(X, Z_2) + Y_1 \, \left[\pi_2^D(X, Z_2) - \pi_2^M(X, Z_2)\right] - \varepsilon_2
\end{aligned}$$

For every firm i, $Y_i \in \{0,1\}$ represents the market entry decision of firm i. $\pi_i^M(.)$ and $\pi_i^D(.)$ are functions that represent the profit of firm i under monopoly and under duopoly, respectively. X, Z_1, and Z_2 are exogenous variables which are observable to the researcher and common knowledge to the players. For each player, ε_i is a random variable that represents a component of the fixed cost of player i that is private information of this player. We assume that ε_1 and ε_2 are independent of (X, Z_1, Z_2) and independently distributed between them with standard Normal distributions. Define the CCP functions $P_i(x, z_1, z_2) \equiv \Pr(Y_i = 1 \mid X = z, \, Z_1 = z_1, \, Z_2 = z_2)$ for $i = 1, 2$.

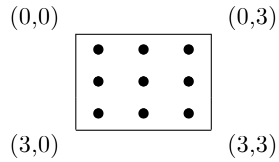
Question 4.1 [10 points]. Describe the equilibrium mapping in the space of CCPs such that a pair of equilibrium probabilities $(P_1(x, z_1, z_2) \, , \, P_2(x, z_1, z_2))$ is a fixed point of that mapping.

Question 4.2 [15 points]. Suppose that: (i) for any value of X, the function $\pi_i^M(X, Z_i)$ depends on Z_i; and (ii) the distribution of (X,Z_1,Z_2) is such that for any value of (X, Z_1) the distribution of Z_2 is non-degenerate, and similarly for any value of (X, Z_2) the distribution of Z_1 is non-degenerate. Prove formally that under conditions (i) and (ii) the payoff functions π_1^M, π_1^D, π_2^M, and π_2^D are nonparametrically identified.

Question 4.3 [5 points]. Discuss the implications of multiple equilibria in the model on the identification and estimation of the payoff function.

PROBLEM 5 [100 points]. Here we consider a **Game of Market Entry and Spatial Location**. The market is a square city where the measure of a side of this square is 3 Km. We represent this city in the two-dimension Euclidean space with vertices at points (0,0), (0,3), (3,0), and (3,3). There are $L = 9$ locations where firms can open stores. The following figure represents the city and the feasible business locations.

Market and feasible business locations (represented with •)



We index locations by ℓ that belongs to the set $\{1, 2, ..., L\}$. There are two potential entrants in the market that we represent as firm i and firm j. Each potential entrant decides whether to operate a store in the market and the location of the store. Let a_i represent the decision of firm/potential entrant i, such that $a_i \in \{0, 1, ..., L\}$ and $a_i = 0$ represents "no entry", and $a_i = \ell > 0$ represents entry in location ℓ. The profit of not being active in the market is normalized to zero. The profit of a store in location ℓ is:

$$\Pi_{i\ell} = POP_\ell \left[\delta_{0i} - \delta_{1i}\ 1\{a_j = \ell\} - \delta_{2i} \left(\sum_{\ell' \in b(\ell)} 1\{a_j = \ell'\} \right) \right] - \alpha_i\ RENT_\ell - \varepsilon_{i\ell}$$

POP_ℓ and $RENT_\ell$ are exogenous variables that represent the population and the average rental price in location ℓ, respectively. a_j represents the entry decision of the competing firm $j \neq i$. $1\{.\}$ is the indicator function such that $1\{a_j = \ell\}$ is the indicator of the event "firm j has decided to have a store in location ℓ. $b(\ell)$ represents the set of locations sharing a boundary with location ℓ. The term $\delta_{0i} - \delta_{1i}\ 1\{a_j = \ell\} - \delta_{2i}$ $\sum_{\ell' \in b(\ell)} 1\{a_j = \ell'\}$ is the variable profit per-potential-customer for firm i, where δ_{0i}, δ_{1i}, and δ_{2i} are parameters that capture the effect of competition. α_i is also a parameter. Finally, $\varepsilon_i = \{\varepsilon_{i\ell} : \ell = 0, 1, ..., L\}$ is a vector of private information variables of firm i at every possible location and it is i.i.d. over firms and locations with a type 1 extreme value distribution.

Given the vector of structural parameters of the model, $\theta \equiv (\alpha_i,\ \delta_{0i},\ \delta_{1i},\ \delta_{2i},\ \alpha_j,\ \delta_{0j},\ \delta_{1j},\ \delta_{2j})'$, and the "landscape" of the exogenous variables over the city locations, $\mathbf{X} \equiv \{POP_\ell,\ RENT_\ell : \ell = 1, 2, ..., L\}$, let $P_{i\ell}(\mathbf{X}, \theta)$ be the probability that firm i enters in location ℓ, i.e., $P_{i\ell}(\mathbf{X}, \theta) = \Pr(a_{i\ell} = 1 \mid \mathbf{X}, \theta)$. And let $\mathbf{P}_i(\mathbf{X}, \theta)$ be the "landscape" of entry probabilities over the L city locations for firm i, i.e., $\mathbf{P}(\mathbf{X}, \theta) \equiv \{P_{i\ell}(\mathbf{X}, \theta) : \ell = 1, 2, ..., L\}$. Given $(\mathbf{X}, \theta)$, the pair of vectors of probabilities $\mathbf{P}_i(\mathbf{X}, \theta)$ and $\mathbf{P}_j(\mathbf{X}, \theta)$ can be defined as a Bayesian Nash Equilibrium of this model.

Question 5.1 [15 points]. Obtain the expression for the expected profit of a potential entrant in location ℓ, obtain the best response probability for entry in location ℓ, and the equilibrium mapping in probability space. Describe $\{\mathbf{P}_i(\mathbf{X}, \theta), \mathbf{P}_j(\mathbf{X}, \theta)\}$ as a fixed point of this equilibrium mapping.

Question 5.2 [10 points]. Suppose that we have cross-sectional data from M cities. For each city, we distinguish 9 geographic regions as in the figure above. Suppose that we observe the entry and location decisions of firms i and j in these M cities:

$$\text{Data} = \{\ POP_{m\ell},\ RENT_{m\ell},\ a_{im},\ a_{jm} : \ell = 0, 1, ..., L;\ m = 1, 2, ..., M\}$$

where we index cities with m. Obtain the expression of the likelihood function (or correspondence) for this model and data.

Question 5.3 [15 points]. Suppose that we treat firms' beliefs about the probabilities of entry of the other firm as incidental parameters. Let the vector of probabilities $\mathbf{B}_i(\mathbf{X}_m) \equiv \{B_{i\ell}(\mathbf{X}_m) : \ell = 0, 1, ..., L\}$ represent firm i' beliefs about the probability of entry of firm j at the different locations of city m. Treating $\mathbf{B}_i(\mathbf{X}_m)$ and $\mathbf{B}_j(\mathbf{X}_m)$ as a vectors of parameters, obtain the expression for the (pseudo) likelihood function $Q(\theta, \mathbf{B}_i, \mathbf{B}_j)$ for the data and model where the choice probabilities in this likelihood are best responses to the beliefs $(\mathbf{B}_i, \mathbf{B}_j)$.

Question 5.4 [10 points]. Show that under the assumption of rational beliefs, we can obtain Nonparametric Reduced Form estimates of firms' beliefs $\mathbf{B}$. Given this consistent estimator of beliefs, propose a two-step consistent estimator of the vector of structural parameters θ.

Question 5.5 [50 points]. The STATA datafile `eco2901_problemset_01_2013.dta` contains a cross-sectional dataset as the one described in Question 5.2 for $M = 1,000$ cities or metropolitan areas.

(a) Use these data to obtain a reduced for estimator of the CCPs $\{P_{i\ell}\}$ using a McFadden's Conditional Logit model.

(b) Using the reduced form estimates in (a), obtain a two-step estimator of the vector of structural parameters θ.

(c) Interpret the results.

Final Exam 2013

INDUSTRIAL ORGANIZATION II (EC 2901)

University of Toronto. Department of Economics. Winter 2013

Instructor: Victor Aguirregabiria

FINAL EXAM: April 11, 2013. From 9:00-12:00 (3 hours)

INSTRUCTIONS: The exam consists of two sets of questions on two papers. Please, try to answer all the questions. Try to allocate your time, and the space of your answers, proportionally to the value of the question (180 points in total $\Leftrightarrow$ 180 minutes). This is a closed book exam. No study aids are allowed.

TOTAL MARKS = 180

PROBLEM 1 (90 points). Answer the following questions on the article *"The Costs of Environmental Regulation in a Concentrated Industry,"* by Stephen Ryan (*Econometrica*, 2012).

Question 1.1 (15 points). Describe the empirical question(s) of the paper and the key features of the empirical strategy that Ryan uses to answer these question. In your opinion, why firm competition and market structure matter to answer these questions?

Question 1.2 (15 points). Describe the components and features of the structural model that the author proposes and estimates. Please, provide a formal answer including the equations and assumptions that describe the different components of the model. What are the main characteristics of the industry that are incorporated into the model? Are important features of the industry missing or ignored? In your opinion, what features of the model are particularly important to answer the empirical questions in the paper?

Question 1.3 (15 points). Describe the different parts in the estimation of the model and in the evaluation of the policy effects. Please, provide a formal answer with the equations that describe the estimated equations, assumptions, and methods. What are the most important econometric problems / challenges that this estimation should deal with?

Question 1.4 (15 points). Provide a critical assessment of the paper. In your opinion, what are the main contributions and limitations of the paper? Try to provide suggestions to improve those limitations.

PROBLEM 2 (90 points). Answer the same questions but on the article *"Dynamic Product Positioning in Differentiated Product Industries: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry,"* by Andrew Sweeting (*Econometrica*, forthcoming 2013). For the sake of completeness, I include the same questions here.

Question 2.1 (15 points). Describe the empirical question(s) of the paper and the key features of the empirical strategy that Sweeting uses to answer these question. In your opinion, why firm competition and market structure matter to answer these questions?

Question 2.2 (15 points). Describe the components and features of the structural model that the author proposes and estimates. Please, provide a formal answer including the equations and assumptions that describe the different components of the model. What are the main characteristics of the industry that are incorporated into the model? Are important features of the industry missing or ignored? In your opinion, what features of the model are particularly important to answer the empirical questions in the paper?

Question 2.3 (15 points). Describe the different parts in the estimation of the model and in the evaluation of the policy effects. Please, provide a formal answer with the equations that describe the estimated equations, assumptions, and methods. What are the most important econometric problems / challenges that this estimation should deal with?

Question 2.4 (15 points). Provide a critical assessment of the paper. In your opinion, what are the main contributions and limitations of the paper? Try to provide suggestions to improve those limitations.

Problem set 2014

Industrial Organization II (ECO 2901)
Winter 2014. Victor Aguirregabiria

Problem Set #1
Due of Friday, February 28, 2014

PROBLEM 1 [100 points]. Competition upstream (manufacturers) and downstream (retailers), and retailers assortment decisions.
There are J brands or product varieties, indexed by $j \in \{1, 2, ..., J\}$. There are F manufacturing firms in the industry that we index by $f \in \{1, 2, ..., F\}$. Manufacturer f produces a subset $\mathcal{J}_f$ of the products, such that the subsets $\{\mathcal{J}_1, \mathcal{J}_2, ..., \mathcal{J}_F\}$ represent a partition of the set $\{1, 2, ..., J\}$. Manufacturers do not sell their products directly to consumers but to retail firms. There is a retail market with L retail firms, indexed by $\ell \in \{1, 2, ..., L\}$, where retailers sell these products to consumers. Retailer ℓ sells an assortment of products $\boldsymbol{a}_\ell \equiv \{a_{1\ell}, a_{2\ell}, ..., a_{J\ell}\}$, where $a_{j\ell} \in \{0, 1\}$ is the indicator of the event "retailer ℓ sells product j". \

The profit of manufacturer f is:

$$\Pi_f = \sum_{j \in \mathcal{J}_f} \left[\sum_{\ell=1}^{L} (w_{j\ell} - c_j) \ q_{j\ell} \right]$$

where: $w_{j\ell}$ is the wholesale price or manufacturer price of product j for retailer ℓ; c_j is the unit cost of producing j; and $q_{j\ell}$ is the quantity of product j purchased by retailer ℓ. The profit of retailer ℓ is:

$$\pi_\ell = \sum_{j=1}^{J} a_{j\ell} \left[(p_{j\ell} - w_{j\ell} - r_\ell - \tau_j) \ q_{j\ell} - K_\ell \right]$$

where: $p_{j\ell}$ is the consumer price of product j charge by retailer ℓ; r_ℓ is the per-unit management cost or retailing cost of retailer retailer ℓ; τ_j is the transportation cost from the production site of product j to the retail market; and K_ℓ is a fixed cost per-product for the retailer.

There are H consumers in the market. Each consumer buys at most one unit of the differentiated product, and has to decide the variety to purchase and the store, (j, ℓ). The demand system is described by the following equations:

$$q_{j\ell} = H \ s_{j\ell} = H \frac{a_{j\ell} \ \exp\{\beta_j + \gamma_\ell - \alpha \ p_{j\ell}\}}{1 + \sum_{\ell'=1}^{L} \sum_{i=1}^{J} a_{i\ell'} \ \exp\{\beta_i + \gamma_{\ell'} - \alpha \ p_{i\ell'}\}}$$

where α, β's, and γ's are parameters. α represents the marginal utility of money. And β_j and γ_ℓ represents consumers' average valuation of product j and store ℓ, respectively. Given this demand system, and taking as given wholesale prices $\{w_{j\ell}\}$, retail firms compete in prices a la Nash-Bertrand. Given the equilibrium in the retail market, manufacturers compete in prices a la Nash-Bertrand. In the simpler version of the model, we assume that product assortments $\boldsymbol{a}_\ell$ are exogenously given. Later, we relax this assumption and consider also retailers' competition in product assortment.

Consider a version of the model where $K_\ell = 0$ and retailers' product assortments $\{\boldsymbol{a}_\ell\}$ are exogenously given.

Question 1.1 [5 points] Derive the system of best response equations that characterizes the Nash-Bertrand equilibrium in the retail market.

Question 1.2 [5 points] Define and describe the demand functions $q_{j\ell} = d_{j\ell}(\mathbf{w})$ from retailers to manufacturers, where $\mathbf{w} = \{w_{j\ell} : j = 1, 2, ..., J, \ \ell = 1, 2, ..., L\}$.

Question 1.3 [5 points] Given the demand systems $q_{j\ell} = d_{j\ell}(\mathbf{w})$, derive the system of best response equations that characterizes the Nash-Bertrand equilibrium in the manufacturer market.

Question 1.4 [5 points] Suppose that manufacturers cannot price discriminate retailers such that $w_{j\ell} = w_j$ for any retailer j. Derive the system of best response equations that characterizes the Nash-Bertrand equilibrium in the manufacturer market without price discrimination.

Question 1.5 [5 points] In the context of this model, explain what is the double marginalization problem.

Now, consider the model where $K_\ell > 0$ and retailers' choose endogenously their product assortments $\boldsymbol{a}_\ell$. Suppose that retailers play a two-stage game where they first choose their product assortments $\{a_\ell\}$, and then they compete in prices taking these assortments as given. The second stage of this game is simply the Nash-Bertrand game in Question 1.1 above. Let $VP_\ell(\boldsymbol{a}_\ell, \boldsymbol{a}_{-\ell})$ be the Nash-Bertrand equilibrium **variable profit** for retailer ℓ provided that he has assortment $\boldsymbol{a}_\ell$ and the other retailers in the market have assortments $\boldsymbol{a}_{-\ell} = \{\boldsymbol{a}_{\ell'} : \ell' \neq \ell\}$. In the first stage of the game (assortment choice), retailers compete a la Nash.

Question 1.6 [5 points] Describe the Nash equilibrium of assortment choice.

Question 1.7 [10 points] Consider a simplified version of the model with: two manufacturers ($F = 2$); two products ($J = 2$) where $j = 1$ is produced by firm 1, and $j = 2$ is produced by firm 2; and two retail firms, $\ell = 1$ and $\ell = 2$. Retailer 1's assortment is fixed and it consists of only product 1, i.e., $a_{11} = 1$ and $a_{21} = 0$. Retailer 2's assortment is endogenous and it is based on the profit maximization of this retailer. (a) Describe in detail the equilibrium in this model. (b) Provide conditions under which retailer 2 decides to sell only product 2. (c) Is it possible to have an equilibrium where both retailers sell only product 1? Explain.

Consider the simplified model in Question 1.7. There are M separate retail markets, that we index by m, where M is large. The two retailers are active in all these markets (we ignore here market entry decisions). Competition between retailers occurs at the local market level. However, competition between manufacturers takes place at the national level such that wholesale prices $\{w_{11}, w_{12}, w_{22}\}$ are the same at every local market m. The variables τ_{1m} and τ_{2m} represent the unit transportation costs from the production sites of products 1 and 2, respectively, to local market m. In each local market, consumer demand follows the structure described above but with two new features: (1) market size H_m varies exogenously across markets; and (2) consumer average willingness to pay for product-retailer (j, ℓ) in market m is equal to $\beta_j + \gamma_\ell + \xi_{j\ell m}$ where β's and γ's are the parameters that we have described above, and $\xi_{j\ell m}$ is a random variable that is unobservable to the researcher and it is i.i.d. across markets with zero mean.

Suppose that the researcher has data from this industry. The dataset includes the following information for each local market m: (a) the assortment choice of firm 2, $(a_{12m}, a_{22m}) \in \{(1,0), (0,1), (1,1)\}$; (b) retail prices and quantities, $(p_{j\ell m}, q_{j\ell m})$, for every active product-retailer (j, ℓ); (c) market size H_m; and (d) unit transportation costs τ_{1m} and τ_{2m}. The dataset includes also wholesale prices and quantities at the national level, $\{w_{j\ell}, Q_{j\ell} : (j, \ell) = (1,1), (1,2), (2,2)\}$ where $Q_{j\ell} \equiv \sum_{m=1}^{M} q_{j\ell m}$.

The researcher is interested in estimating the vector of structural parameters of the model, θ, that includes: the average quality of the two products $\{\beta_1, \beta_2\}$; the average quality of the two retailers $\{\gamma_1, \gamma_2\}$; the demand-price sensitivity parameter α; the unit retail costs $\{r_1, r_2\}$; the fixed cost of retailer 2, K_2; and the unit manufacturing costs $\{c_1, c_2\}$.

Question 1.8 [5 points] Write the consumer demand system as a linear regression-like system of equations.

Question 1.9 [10 points] Discuss the endogeneity problems associated to the estimation of parameters $\{\beta_1, \beta_2, \gamma_1, \gamma_2\}$ in this system of linear regression equations. In particular, explain the endogeneity (selection) problem related to the assortment choice of retailer 2.

Question 1.10 [10 points] Propose a method to estimate consistently the demand parameters $\{\beta_1, \beta_2, \gamma_1, \gamma_2\}$ in this model. Explain the necessary condition for identification, and the different steps in the implementation of this method.

Question 1.11 [10 points] Given consistent estimates of $\{\beta_1, \beta_2, \gamma_1, \gamma_2\}$, consider the estimation of the unit retail costs $\{r_1, r_2\}$ from the optimal pricing equations. (a) Write the expression for these equations. (b) Is there a selection problem in the estimation of this equation? Why/why not? (c) Propose a method to estimate consistently the parameters $\{r_1, r_2\}$.

Question 1.12 [10 points] Given consistent estimates of $\{\beta_1, \beta_2, \gamma_1, \gamma_2, r_1, r_2\}$, consider the estimation of the fixed cost K_2 from the optimal assortment decision of retailer 2, $(a_{12m}, a_{22m}) \in \{(1,0), (0,1), (1,1)\}$.

(a) Write the discrete choice econometric model for this estimation. (b) Propose a method to estimate consistently the parameter K_2 in this model.

Question 1.13 [5 points] Given consistent estimates of $\{\beta_1, \beta_2, \gamma_1, \gamma_2, r_1, r_2, K_2\}$, consider the estimation of the unit production cost parameters $\{c_1, c_2\}$ from the best response pricing decisions of manufacturers. Explain how to obtain consistent estimates of these parameters.

Question 1.14 [10 points] The model has been estimated with data from an industry where manufacturing firms can price discriminate retailers. Suppose that the researcher is interested in the following empirical question: what would be the level of prices, market shares, firms' profits, and consumer welfare in the different local markets if price discrimination were illegal? Describe a counterfactual experiment that answers this empirical question. Explain how to implement this experiment to obtain counterfactual values of all the endogenous variables.

PROBLEM 2 [20 points]. Conjectural variations with differentiated product.
We have an industry with J manufacturing firms where each firm produces a variety of a differentiated product. Suppose that firms compete in prices but firms' beliefs about other firms actions do not correspond to the Nash axiom. Instead, any firm j beliefs that for any $k \neq j$, $\dfrac{\partial p_k}{\partial p_j} = \theta$, where θ is a conjectural variation parameter.

Question 2.1 [5 points] Describe an equilibrium of this model under the conjectural variation assumption of $\dfrac{\partial p_k}{\partial p_j} = \theta$.

Suppose that the industry is such that J is large. The researcher has data on $\{q_j, p_j, X_j : j = 1, 2, ..., J\}$. The researcher has estimated the demand system such that the market share function $\sigma_j(\mathbf{p}, \mathbf{X}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\beta}})$ are known, where $\widehat{\boldsymbol{\xi}}$ and $\widehat{\boldsymbol{\beta}}$ represent consistent estimates of unobserved product characteristics and demand parameters, respectively. The marginal cost for product j is $MC_j = X_j \gamma + \omega_j$, where γ is a vector of parameters and ω_j represents the effect of unobserved characteristics on marginal costs.

Question 2.2 [5 points] Present a regression-like equation for the estimation of the parameters γ and θ. Explain the problems for the estimation of these parameters.

Question 2.3 [10 points] Propose a method for the consistent estimation of γ and θ. Explain the necessary conditions for identification/consistency.

PROBLEM 3 [80 points].
The STATA datafile `eco2901_problemset_01_2014_airlines_data.dta` 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 3.1 [20 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 hub size at the origin airport (in dollars per million people)? 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 3.2 [20 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 3.1.

Question 3.3 [20 points]: Consider the Nested Logit model in Question 3.2. 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 3.2.

Question 3.4 [20 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.

Other Problems

A. STATIC (MYOPIC) ENTRY-EXIT GAME

We first consider a static (not forward-looking) version of the entry-exit game. A Bayesian Nash Equilibrium (BNE) in this game can be described as a pair of probabilities, $\{P_{MD}(x_t), P_{BK}(x_t)\}$ solving the following system of equations:

$$\begin{aligned} P_{MD}(x_t) &= \Phi\left(Z^P_{MDt}\ \theta_{MD}\right) \\ P_{BK}(x_t) &= \Phi\left(Z^P_{BKt}\ \theta_{BK}\right) \end{aligned}$$

where $\Phi(.)$ is the CDF of the standard normal.

QUESTION 2. [10 POINTS] For every possible value of the state x_t (i.e., 24 values) obtain all the BNE of the static entry game.

Hint: Define the functions $f_{MD}(P) \equiv \Phi\left(Z^P_{MDt}\ \theta_{MD}\right)$ and $f_{BK}(P) \equiv \Phi\left(Z^P_{BKt}\ \theta_{BK}\right)$. Define also the function $g(P) \equiv P - f_{MD}(f_{BK}(P))$. A BNE is zero of the function $g(P)$. You can search for all the zeroes of $g(P)$ in different ways, but in this case the simpler method is to consider a discrete grid for P in the interval $[0,1]$, e.g., uniform grid with 101 points.

For some values of the state vector x_t, the static model has multiple equilibria. To answer Questions 3 to 5, assume that, in the population under study, the "equilibrium selection mechanism" always selects the equilibrium with the higher probability that MD is active in the market.

Let X be the set of possible values of x_t. And let $\mathbf{P}^0 \equiv \{P^0_{MD}(x), P^0_{BK}(x) : x \in X\}$ be the equilibrium probabilities in the population. Given $\mathbf{P}^0$ and the transition probability matrix for market size, F_S. We can obtain the steady-state distribution of x_t. Let $p^*(x_t)$ be the steady-state distribution. By definition, for any $x_{t+1} \in X$:

$$\begin{aligned} p^*(x_{t+1}) &= \sum_{x_t \in X} p^*(x_t)\ \Pr(x_{t+1}|x_t) \\ &= \sum_{x_t \in X} p^*(x_t)\ F_S(S_{t+1}|S_t) \\ &\quad \left[P^0_{MD}(x_t)\right]^{a_{MDt+1}} \left[1 - P^0_{MD}(x_t)\right]^{1-a_{MDt+1}} \left[P^0_{BK}(x_t)\right]^{a_{BKt+1}} \left[1 - P^0_{BK}(x_t)\right]^{1-a_{BKt+1}} \end{aligned}$$

QUESTION 3. [10 POINTS] Compute the steady-state distribution of x_t in the population.

QUESTION 4. [20 POINTS] Using the values of P^0, F_S and p^* obtained above, simulate a data set $\{x_{mt} : t = 0, 1, ..., T; m = 1, 2, ..., M\}$ for $M = 500$ local markets and $T + 1 = 6$ years with the following features: (1) local markets are independent; and (2) the initial states x_{m0} are random draws from the steady-state distribution p^*. Present a table with the mean values of the state variables in x_t and with the sample frequencies for the following events: (1) MD is a monopolist; (2) BK is a monopolist; (3) duopoly; (4) MD is active given that (conditional) he was a monopolist at the beginning of the year (the same for BK); (5) MD is active given that BK was a monopolist at the beginning of the year (the same for BK); (6) MD is active given that there was a duopoly at the beginning of the year (the same for BK); and (7) MD is active given that there were no firms active at the beginning of the year (the same for BK).

QUESTION 5. [20 POINTS] Use the simulated data in Question 4 to estimate the structural parameters of the model. Implement the following estimators: (1) two-step PML using a frequency estimator of $\mathbf{P}^0$ in the first step; (2) two-step PML using random draws from a U(0,1) for $\mathbf{P}^0$ in the first step; (3) 20-step PML using a frequency estimator of $\mathbf{P}^0$ in the first step; (4) 20-step PML using random draws from a U(0,1) for P^0 in the first step; and (5) NPL estimator based on 10 NPL fixed points (i.e., 10 different initial $P's$). Comment the results.

QUESTION 6. [30 POINTS] Suppose that the researcher knows that local markets are heterogeneous in their market size, but he does not observed market size S_{mt}. Suppose that the researcher assumes that market size is constant over time but it varies across markets, and it has a uniform distribution with discrete support $\{4, 5, 6, 7, 8, 9\}$. Obtain the NPL estimator under this assumption (use 20 NPL fixed points). Comment the results.

QUESTION 7. [30 POINTS] Use the previous model (both the true model and the model estimated in Question 5) to evaluate the effects of a value added tax. The value added tax is paid by the retailer and it is such that the parameters θ_i^M and θ_i^D are reduced by 10%. Obtain the effects of this tax on average firms' profits, and on the probability distribution of market structure.

B. DYNAMIC ENTRY-EXIT GAME

Now, consider the dynamic (forward-looking) version of the entry-exit game. A Markov Perfect Equilibrium (MPE) in this game can be described as a vector of probabilities $\mathbf{P} \equiv \{P_i(x_t) : i \in \{MD, BK\}, x_t \in X\}$ such that, for every (i, x_t):

$$P_i(x_t) = \Phi\left(\tilde{Z}_{it}^P\ \theta_{MD} + \tilde{e}_{it}^P\right)$$

where $\tilde{Z}_{it}^P$ and $\tilde{e}_{it}^P$ are defined in the class notes.

QUESTION 8. [20 POINTS] Obtain the MPE that we obtain when we iterate in the equilibrium mapping starting with an initial $\mathbf{P} = \mathbf{0}$. Find other MPEs.

QUESTION 9. [10 POINTS] Compute the steady-state distribution of x_t in the population.

QUESTION 10. [20 POINTS] The same as in Question 4 but using the dynamic game and the MPE in Question 8.

QUESTION 11. [20 POINTS] The same as in Question 5 but using the dynamic game and the MPE in Question 8.

QUESTION 12. [30 POINTS] The same as in Question 6 but using the dynamic game and the MPE in Question 8.

QUESTION 13. [30 POINTS] The same as in Question 7 but using the dynamic game and the MPE in Question 8.

PROBLEM B

Consider a market with N firms who can potentially operate in it. We index firms by $i \in \{1, 2, ..., N\}$. Firms produce and sell a differentiated product. There are S consumers and each consumer buys at most one unit (per period) of this differentiated product. A consumer (indirect) utility of buying firm $i's$ product is:

$$U_i = w_i - p_i + \varepsilon_i$$

w_i is the "quality" of product i which is valued in the same way by every consumer. p_i is the price. And $\{\varepsilon_1, \varepsilon_2, ..., \varepsilon_N\}$ are consumer specific preferences which are i.i.d. with a type 1 extreme value distribution with dispersion parameter α. The utility of not buying any of these products is normalized to zero. For simplicity, we consider that there are only two levels of quality, high and low: $w_i \in \{w_L, w_H\}$, with $w_L < w_H$. Firms choose endogenously their qualities and prices. They also decide whether to operate in the market or not. Let n_L and n_H be the number of active firms with low and high quality products, respectively. Then, the demand for an active firm with quality w_i and price p_i is:

$$q_i = \frac{S \ \exp\left\{\dfrac{w_i - p_i}{\alpha}\right\}}{1 + n_L \ \exp\left\{\dfrac{w_L - p_L}{\alpha}\right\} + n_H \ \exp\left\{\dfrac{w_H - p_H}{\alpha}\right\}}$$

where we have imposed the (symmetric) equilibrium restriction that firms with the same quality charge the same price. Inactive firms get zero profit. The profit of an active firm is:

$$\Pi_i = (p_i - c(w_i))\, q_i - F(w_i)$$

where $c(w_i)$ and $F(w_i)$ are the (constant) marginal cost and the fixed cost of producing a product with quality w_i. Each firm decides: (1) whether or not to operate in the market; (2) the quality of its product; and (3) its price. The game that firms play is a sequential game with the following two steps:

Step 1: Firms make entry and quality decisions. This determines n_L and n_H.
Step 2: Given (n_L, n_H), firms compete in prices a la Bertrand.

We start describing the Bertrand equilibrium at step 2 of the game.

QUESTION 1. [10 POINTS] Show that the best response functions of the Bertrand game in step 2 have the following form.

$$
\begin{aligned}
p_L &= c_L + \alpha \left[1 - \frac{\exp\left\{\dfrac{w_L - p_L}{\alpha}\right\}}{1 + n_L \ \exp\left\{\dfrac{w_L - p_L}{\alpha}\right\} + n_H \ \exp\left\{\dfrac{w_H - p_H}{\alpha}\right\}} \right]^{-1} \\
p_H &= c_H + \alpha \left[1 - \frac{\exp\left\{\dfrac{w_H - p_H}{\alpha}\right\}}{1 + n_L \ \exp\left\{\dfrac{w_L - p_L}{\alpha}\right\} + n_H \ \exp\left\{\dfrac{w_H - p_H}{\alpha}\right\}} \right]^{-1}
\end{aligned}
$$

ANSWER:

Note that equilibrium prices depend on (n_L, n_H).

QUESTION 2. [30 POINTS] Write a computer program that computes equilibrium prices in this Bertrand game. For given values of the structural parameters (e.g., $\alpha = 1$**,** $w_L = 2$**,** $w_H = 4$**,** $c_L = 1$**,** $c_H = 2$**) calculate equilibrium prices for every possible combination of** (n_L, n_H) **given that** $N = 4$**. Present the results in a table.**

n_L	n_H	p_L	p_H
1	0	?	?
0	1	?	?
1	1	?	?
2	0	?	?
...	...	...	...

Now, consider the game at step 1. It is useful to define the indirect variable profit function that results from the Bertrand equilibrium in step 2 of the game. Let $\Pi_L(n_L, n_H)$ and $\Pi_H(n_L, n_H)$ be this indirect variable profit, i.e., $\Pi_L(n_L, n_H) = (p_L - c_L)q_L$ and $\Pi_H(n_L, n_H) = (p_H - c_H)q_H$, where prices and quantities are equilibrium ones.

QUESTION 3. [10 POINTS] Show that: $\Pi_L(n_L, n_H) = \alpha S q_L/(S - q_L)$ and $\Pi_H(n_L, n_H) = \alpha S q_H/(S - q_H)$.

Let $n_{L(-i)}$ and $n_{H(-i)}$ be the number of low and high quality firms, respectively, excluding firm i. Let's use $w_i = \varnothing$ to represent no entry. And let $b(n_{L(-i)}, n_{H(-i)})$ be the best response mapping of a firm at step 1 of the game.

QUESTION 4. [10 POINTS] Show that the best response function $b(n_{L(-i)}, n_{H(-i)})$ **can be described as follows:**

$$
b(n_{L(-i)}, n_{H(-i)}) = \begin{cases}
\varnothing \quad if & \left[\begin{array}{c} \{\Pi_L(n_{L(-i)}+1, n_{H(-i)}) - F_L < 0\} \\ \text{and } \{\Pi_H(n_{L(-i)}, n_{H(-i)}+1) - F_H < 0\} \end{array} \right] \\[2ex]
w_L \quad if & \left[\begin{array}{c} \{\Pi_L(n_{L(-i)}+1, n_{H(-i)}) - F_L \geq 0\} \\ \text{and } \{\Pi_L(n_{L(-i)}+1, n_{H(-i)}) - F_L > \Pi_H(n_{L(-i)}, n_{H(-i)}+1) - F_H\} \end{array} \right] \\[2ex]
w_H \quad if & \left[\begin{array}{c} \{\Pi_H(n_{L(-i)}, n_{H(-i)}+1) - F_H \geq 0\} \\ \text{and } \{\Pi_L(n_{L(-i)}+1, n_{H(-i)}) - F_L \leq \Pi_H(n_{L(-i)}, n_{H(-i)}+1) - F_H\} \end{array} \right]
\end{cases}
$$

Now, suppose that a component of the fixed cost is private information of the firm: i.e., $F_i(w_L) = F_L + \xi_{iL}$ and $F_i(w_H) = F_H + \xi_{iH}$, where F_L and F_H are parameters and ξ_{iL} and ξ_{iH} are private information variables which are iid extreme value distributed across firms. In this Bayesian game a firm's strategy is a function of his own private information $\xi_i \equiv (\xi_{iL}, \xi_{iH})$ and of the common knowledge variables (i.e., parameters of the model and market size S). Let $\omega(\xi_i, S)$ be a firm's strategy function. A firm's strategy can be also described in terms of two probabilities: $P_L(S)$ and $P_H(S)$, such that:

$$P_L(S) \equiv \int I\{\omega(\xi_i, S) = w_L\} \; dF_\xi(\xi_i)$$

$$P_H(S) \equiv \int I\{\omega(\xi_i, S) = w_H\} \; dF_\xi(\xi_i)$$

where $I\{.\}$ is the indicator function and $F_\xi(.)$ is the CDF of ξ_i.

QUESTION 5. [20 POINTS] Show that a Bayesian Nash Equilibrium (BNE) in this game is a pair (P_L, P_H) that is a solution to the following fixed problem:

$$P_L = \frac{\exp\{\Pi_L^e(P_L, P_H) - F_L\}}{1 + \exp\{\Pi_L^e(P_L, P_H) - F_L\} + \exp\{\Pi_H^e(P_L, P_H) - F_H\}}$$

$$p_H = \frac{\exp\{\Pi_H^e(P_L, P_H) - F_H\}}{1 + \exp\{\Pi_L^e(P_L, P_H) - F_L\} + \exp\{\Pi_H^e(P_L, P_H) - F_H\}}$$

with:

$$\Pi_L^e(P_L, P_H) = \sum\nolimits_{n_{L(-i)}, n_{H(-i)}} \Pi_L(n_{L(-i)} + 1, n_{H(-i)}) \; T(n_{L(-i)}, n_{H(-i)} | N-1, P_L, P_H)$$

$$\Pi_H^e(P_L, P_H) = \sum\nolimits_{n_{L(-i)}, n_{H(-i)}} \Pi_H(n_{L(-i)}, n_{H(-i)} + 1) \; T(n_{L(-i)}, n_{H(-i)} | N-1, P_L, P_H)$$

where $T(x, y|n, p_1, p_2)$ is the PDF of a trinomial distribution with parameters (n, p_1, p_2).

QUESTION 6. [50 POINTS] Write a computer program that computes the BNE in this entry/quality game. Consider $N = 4$. For given values of the structural parameters, calculate the equilibrium probabilities $(P_L(S), P_H(S))$ for a grid of points for market size S. Present a graph for $(P_L(S), P_H(S))$ (on the vertical axis) on S (in the horizontal axis). Does the proportion of high quality firms depend on market size?

QUESTION 7. [30 POINTS] Define the function $\lambda(S) \equiv P_H(S)/P_L(S)$ that represents the average ratio between high and low quality firms in the market. Repeat the same exercise as in Question 1.6. but for three different values of the ratio F_H/F_L. Present a graph of $\lambda(S)$ on S for the three values of F_H/F_L. Comment the results.

QUESTION 8. [50 POINTS] A regulator is considering a policy to encourage the production of high quality products. The policy would provide a subsidy of 20% of the additional fixed cost of producing a high quality product. That is, the new fixed cost of producing a high quality product would be $F_H^* = F_H - 0.20 * (F_H - F_L)$. Given a parametrization of the model, obtain the equilibrium before and after the policy and calculate the effect of the policy on: (1) prices; (2) quantities; (3) firms' profits; (4) average consumers' surplus; and (5) total surplus.

Suppose that the researcher observes a random sample of M isolated markets, indexed by m, where these N firms compete. More specifically, the researcher observes:

$$Data = \{S_m,\ n_{Hm},\ n_{Lm},\ q_{Hm},\ q_{Lm}, p_{Hm}, p_{Lm} : m = 1, 2, ..., M\}$$

For instance, consider data from the hotel industry in a region where "high quality" is defined as four stars or more (low quality as three stars or less). We incorporate two sources of market heterogeneity in the econometric model (i.e., unobservables for the researcher).

(A) Consumers' average valuations: $w_{Lm} = w_L + \eta_m^{(wL)}$ and $w_{Hm} = w_H + \eta_m^{(wH)}$, where $\eta_m^{(wL)}$ and $\eta_m^{(wH)}$ are zero mean random variables.
(B) Marginal costs: $c_{Lm} = c_L + \eta_m^{(cL)}$ and $c_{Hm} = c_H + \eta_m^{(cH)}$, where $\eta_m^{(cL)}$ and $\eta_m^{(cH)}$ are zero mean random variables.

We assume that the vector of unobservables $\eta_m \equiv \{\eta_m^{(wL)}, \eta_m^{(wH)}, \eta_m^{(cL)}, \eta_m^{(cH)}\}$ is iid over markets and independent of market size S_m. We also assume that these variables are common knowledge. We want to use these data to estimate the structural parameters $\theta = \{\alpha, w_j, c_j, F_j : j = L, H\}$.

QUESTION 9. [30 POINTS] Show that the econometric model can be described in terms of three sets of equations.

(1) Demand equations: For $j \in \{L, H\}$ let s_{jm} be the market share q_{jm}/S_m. Then:

$$\ln\left(\frac{s_{jm}}{1 - s_{Lm} - s_{Hm}}\right) = \frac{w_j}{\alpha} - \frac{1}{\alpha}\ p_{jm} + \frac{\eta_m^{(wj)}}{\alpha} \qquad \text{if } n_{jm} > 0$$

(2) Price equations: For $j \in \{L, H\}$:

$$p_{jm} = c_j + \alpha\left(\frac{1}{1 - s_{jm}}\right) + \eta_m^{(cj)} \qquad \text{if } n_{jm} > 0$$

(3) Entry/Quality choice: Suppose that from the estimation of (1) and (2) we can obtain consistent estimates of η_m as residuals. After that estimation, we can treat η_m as "observable" (though we should account for estimation error). Then,

$$\Pr(n_{Lm}, n_{Hm}|S_m, \eta_m) = T(n_{Lm}, n_{Hm}|N, P_L(S_m, \eta_m), P_H(S_m, \eta_m))$$

where $P_L(S_m, \eta_m), P_H(S_m, \eta_m)$ are equilibrium probabilities in market m.

QUESTION 10. [30 POINTS] Discuss in detail the econometric issues in the estimation of the parameters $\{w_L, w_H, \alpha\}$ from the demand equations: for $j \in \{L, H\}$:

$$\ln\left(\frac{s_{jm}}{1 - s_{Lm} - s_{Hm}}\right) = \frac{w_j}{\alpha} - \frac{1}{\alpha}\ p_{jm} + \frac{\eta_m^{(wj)}}{\alpha} \qquad \text{if } n_{jm} > 0$$

Propose and describe in detail a method that provides consistent estimates of $\{w_L, w_H, \alpha\}$.

QUESTION 11. [30 POINTS] Suppose for the moment that α has not been estimated from the demand equations. Discuss in detail the econometric issues in the estimation of the parameters $\{c_L, c_H, \alpha\}$ from the pricing equations: for $j \in \{L, H\}$:

$$p_{jm} = c_j + \alpha\left(\frac{1}{1 - s_{jm}}\right) + \eta_m^{(cj)} \qquad \text{if } n_{jm} > 0$$

Propose and describe in detail a method that provides consistent estimates of $\{c_L, c_H, \alpha\}$. What if α has been estimated in a first step from the demand equations? Which are the advantages of a joint estimation of demand and supply equations?

QUESTION 12. [50 POINTS] For simplicity, suppose that the parameters $\{w_L, w_H, c_L, c_H, \alpha\}$ are known and that η_m is observable (i.e., we ignore estimation error from the first step estimation). We want to estimate the fixed costs F_L and F_H using information on firms' entry/quality choices. Discuss in detail the econometric issues in the estimation of these parameters. Propose and describe in detail a method that provides consistent estimates of $\{F_L, F_H\}$.

QUESTION 13. [50 POINTS] Suppose that you incorporate a third source of market heterogeneity in the model:

(C) Fixed costs: $F_{Lm} = F_L + \eta_m^{(FL)}$ and $F_{Hm} = F_H + \eta_m^{(FH)}$, where $\eta_m^{(FL)}$ and $\eta_m^{(FH)}$ are zero mean random variables, and they are common knowledge to the players.

Explain which are the additional econometric issues in the estimation of $\{F_L, F_H\}$ when we have these additional unobservables. Propose and describe in detail a method that provides consistent estimates of $\{F_L, F_H\}$ and the distribution of $\{\eta_m^{(FL)}, \eta_m^{(FH)}\}$.

QUESTION 14. [50 POINTS] Consider the econometric model without $\{\eta_m^{(FL)}, \eta_m^{(FH)}\}$. Suppose that S_m is log normally distributed and $\eta_m \equiv \{\eta_m^{(wL)}, \eta_m^{(wH)}, \eta_m^{(cL)}, \eta_m^{(cH)}\}$ has a normal distribution with zero means. Generate a random sample of $\{S_m, \eta_m\}$ with sample size of $M = 500$ markets. Given a parametrization of the model, for every value $\{S_m, \eta_m\}$ in the sample, solve the model and obtain the endogenous variables $\{n_{Hm},\ n_{Lm},\ q_{Hm},\ q_{Lm},\ p_{Hm},\ p_{Lm}\}$. Present a table with the summary statistics of these variables: e.g., mean, median, standard deviation, minimum, maximum.

QUESTION 15. [50 POINTS] Write a computer program that implements the method for the estimation of the demand that you proposed in Question 10. Apply this method to the data simulated in Question 14. Present and comment the results.

QUESTION 16. [50 POINTS] Write a computer program that implements the method for the estimation of the pricing equations that you proposed in Question 11. Apply this method to the data simulated in Question 14. Present and comment the results.

QUESTION 17. [100 POINTS] Write a computer program that implements the method for the estimation of the entry/quality choice game that you proposed in Question 12. Apply this method to the data simulated in Question 14. Present and comment the results.

QUESTION 18. [50 POINTS] Use the estimated model to evaluate the policy in question 8. Present a table that compares the average (across markets) "actual" and estimated effects of the policy on: (1) prices; (2) quantities; (3) firms' profits; (4) average consumers' surplus; and (5) total surplus.

PROBLEM C

In the paper *"The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish,"* (REStud, 2000), Angrist, Graddy and Imbens consider the following random coefficients model of supply and demand for an homogeneous product:

$$\begin{aligned} \text{Inverse Demand:} \quad p_t &= x_t\, \beta^D - \left(\alpha^D + \upsilon_t^D\right)\, q_t + \varepsilon_t^D \\ \text{Inverse Supply:} \quad p_t &= x_t\, \beta^S + \left(\alpha^S + \upsilon_t^S\right)\, q_t + \varepsilon_t^S \end{aligned}$$

where p_t is logarithm of price; q_t is the logarithm of the quantity sold; and ε_t^D, ε_t^S, υ_t^D and υ_t^S are unobservables which have zero mean conditional on x_t. The variables υ_t^D and υ_t^S account for random shocks in the price elasticities of demand and supply. Suppose that the researcher has a sample $\{p_t, q_t, x_t : t = 1, 2, ..., n\}$ and is interested in the estimation of the demand parameters β^D and α^D.

1. (a) Explain why instrumental variables (or 2SLS) provides inconsistent estimates of the parameters β^D and α^D.

   (b) Descrine an estimation method that provides consistent estimates of β^D and α^D.

PROBLEM D

Mitsubishi entered the Canadian automobile market in September 2002. You can consider this to be an exogenous change. Subsequently, the firm had to decide in which local markets to open dealerships. This, you should consider to be endogenous choices.

1. (a) How could you use this type of variation to estimate a model of entry like Bresnahan & Reiss (1988, 1990, 1991)? What variation in the data will be useful to identify which underlying economic parameters? How would you learn about or control for the competitiveness of market operation?

   (It is not necessary to derive any equations, although you can if it helps your exposition.).

   (b) Could you use the same data to estimate an entry model like Berry (1992)? How?

   (c) How would you use data for this industry to estimate the lower bound on concentration in the sense of Sutton?

   (d) Give an example of an economic question that you would be able to address with this type of variation over time —entry by a new firm— that the previous authors were unable to address using only cross sectional data.

PROBLEM E

In the paper *"The valuation of new goods under perfect and imperfect competition,"* Jerry Hausman estimates a demand system for ready-to eat cereals using panel data on quantities and prices for multiple markets (cities), brands and quarters. The demand system is (Deaton-Muellbauer demand system):

$$w_{jmt} = \alpha_j^0 + \alpha_m^1 + \alpha_t^2 + \sum_{k=1}^{J} \beta_{jk} \ \ln(p_{kmt}) + \gamma_j \ \ln(x_{mt}) + \varepsilon_{jmt}$$

where: j, m and t are the product, market (city) and quarter subindexes, respectively; x_{mt} represents exogenous market characteristics such as population and average income. There are not observable cost shifters. The terms α_j^0, α_m^1 and α_t^2 represent product, market and time effects, respectively, which are captured using dummies. As instruments for prices, Hausman uses average prices is nearby markets. More specifically, the instrument for price p_{jmt} is z_{jmt} which is defined as:

$$z_{jmt} = \frac{1}{\#\left(R_m\right)} \sum_{\substack{m' \neq m \\ m' \in R_m}} p_{jm't}$$

1. where R_m is the set of markets nearby market m, and, $\#\left(R_m\right)$ is the number of elements in that set.
   (a) Explain under which economic assumptions, on supply or price equations, these instruments are valid.
   (b) Describe how Deaton-Muellbauer demand system can be used to calculate the value of a new product.
   (c) Comment the limitations of this approach as a method to evaluate the effects of new product on consumers' welfare and firms' profits.
   (d) Explain how the empirical literature on demand models in characteristics space deals with some of the limitations that you have mentioned in question (c).

PROBLEM F

Consider Berry-Levinshon-Pakes (BLP) model for the demand of a differentiated product. The (indirect) utility of buying product j for consumer i is:

$$U_{ij} = (\beta_1 + \omega_{1i})x_{1j} + ... + (\beta_K + \omega_{Ki})x_{Kj} - \alpha\ p_j + \xi_j + \varepsilon_{ij}$$

where α, β_1, ..., and β_K are parameters; $\omega_i \equiv (\omega_{1i}, \omega_{2i}, ..., \omega_{Ki})$ is a vector of normal random variables (with zero mean); and $\varepsilon_i \equiv (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iJ})$ is a vector of independent extreme value random variables.

1. (a) Describe in detail BLP estimation method.
   (b) Explain why it is important to allow for consumer heterogeneity in the marginal utility with respect to product characteristics.
   (c) A key identifying assumption in BLP method is that unobserved product characteristics, ξ_j, are not correlated with observed product characteristics other than price, $(x_{1j}, x_{2j}, ..., x_{Kj})$. Comment on this assumption.
   (d) Suppose that there is only one observable product characteristic, x_j, that we can interpret as a measure of product quality. Let x_j^* is the "true" quality of product j, which is unobservable to the researcher. That is, $x_j = x_j^* + e_j$ where e_j is measurement error which is assumed independent of x_j^*. According to this model, the unobservable ξ_j is equal to $-\beta e_j$. Show that the type of instrumental variables proposed by BLP can still be valid in this model with measurement error in quality.

PROBLEM G

Consider an oligopoly industry in which competition takes place at the level of local markets. For concreteness, suppose that there are only two firms in the industry: firm 1 and firm 2. There are M local markets, where M is a large number. Consider the following adaptation to this industry of the simultaneous equations model in *Olley and Pakes (Econometrica, 1996)*.

Production Function: $y_{imt} = \alpha_{Li}\ l_{imt} + \alpha_{Ki}\ k_{imt} + \omega_{imt} + e_{imt}$

Investment Function: $i_{imt} = f_i\left(k_{1mt}, k_{2mt}, \omega_{1mt}, \omega_{2mt}, r_{mt}\right)$

Stay-in-the-market decision: $\chi_{imt} = I\{\omega_{imt} \geq \omega_i^*\left(k_{1mt}, k_{2mt}, r_{mt}\right)\}$

where: i is the firm subindexl; m is the local-market subindex; t is the time subindex; r_{mt} represents input prices in market m at period t; and all the other variables and parameters have the same interpretation as in Olley-Pakes. Following Olley-Pakes we assume that labor is a perfectly flexible input and that new investment is not productivity until next period (i.e., time-to-build). We are interested in the estimation of the production function parameters $\{\alpha_{L1}, \alpha_{K1}, \alpha_{L2}, \alpha_{K2}\}$.

1. (a) Explain why a direct application of Olley-Pakes method to this model will not provide consistent estimates of the parameters of interest.

   (b) Describe how Olley-Pakes method can be adapted/extended to this industry and data to obtain a consistent estimator of $\{\alpha_{L1}, \alpha_{K1}, \alpha_{L2}, \alpha_{K2}\}$.

   (c) Suppose that the average productivity of labor is larger in markets where both firms are active (relative to markets where only one of the two firms is active). Mention different hypotheses that might explain this evidence. Explain how one can use the estimated model to measure the contribution of each of these hypothesis to the observed differential in the average productivity of labor.

PROBLEM H

Consider the following description of a hotel industry. There are N firms/hotel chains in the industry. These firms compete in independent local markets (cities). We index hotel chains by $i \in \{1, 2, ..., N\}$ and local markets by $m \in \{1, 2, ..., N\}$. The product that hotels sell is vertically differentiated. For simplicity, we consider that there are only two levels of quality, high (H) and low (L). At each local market, each firm decides whether or not to operate in the market, the quality of its product, and its price. The game that hotel chains play is a sequential game with the following two steps. Step 1: firms make entry and quality decisions. This step determines the number of low and high quality hotels in the market: n_m^L and n_m^H respectively. Step 2: Given (n_m^L, n_m^H), firms compete in prices a la Bertrand. Associated to the Bertrand equilibrium we can define the (indirect) variable profit functions $V_L(n_m^L, n_m^H, S_m)$ and $V_H(n_m^L, n_m^H, S_m)$: i.e., $V_L(n_m^L, n_m^H, S_m)$ ($V_H(n_m^L, n_m^H, S_m)$) is the variable profit of a low (high) quality hotel in a market with size S_m, with n_m^L low quality hotels and with n_m^H high quality hotels. Total operating costs are: $\Pi_{Lim} = V_L(n_m^L, n_m^H, S_m) - F_L - \varepsilon_{Lim}$ and $\Pi_{Him} = V_H(n_m^L, n_m^H, S_m) - F_H - \varepsilon_{Him}$, where F_L and F_H are the fixed costs for low and high quality firms, respectively, and ε_{Lim} and ε_{Him} are private information shocks which are iid extreme value distributed across firms and markets. A firm's strategy can be described in terms of two probabilities: the probability of being active with low quality, P_L, and the probability of being active and high quality, P_H.

1. (a) Show that a Bayesian Nash Equilibrium (BNE) in this game is a pair (P_L, P_H) that is a solution to the following fixed problem:

$$P_L = \frac{\exp\{V_L^e(P_L, P_H) - F_L\}}{1 + \exp\{V_L^e(P_L, P_H) - F_L\} + \exp\{V_H^e(P_L, P_H) - F_H\}}$$

$$P_H = \frac{\exp\{V_H^e(P_L, P_H) - F_H\}}{1 + \exp\{V_L^e(P_L, P_H) - F_L\} + \exp\{V_H^e(P_L, P_H) - F_H\}}$$

with:

$$V_L^e(P_L, P_H) = \sum\nolimits_{n_{L(-i)}, n_{H(-i)}} V_L(n_{L(-i)} + 1, n_{H(-i)})\; T(n_{L(-i)}, n_{H(-i)} | N-1, P_L, P_H)$$

$$V_H^e(P_L, P_H) = \sum\nolimits_{n_{L(-i)}, n_{H(-i)}} V_H(n_{L(-i)}, n_{H(-i)} + 1)\; T(n_{L(-i)}, n_{H(-i)} | N-1, P_L, P_H)$$

where $T(x, y|n, p_1, p_2)$ is the PDF of a trinomial distribution with parameters (n, p_1, p_2).

(b) Suppose that the indirect profit functions $V_L(n_L, n_H, S)$ and $V_H(n_L, n_H, S)$ are known, i.e., they have been estimated using price an quantity data). The researcher observes the sample $\{n_{Hm}, n_{Lm}, S_m : m = 1, 2, ..., M\}$. We want to estimate the fixed costs F_L and F_H using information on firms' entry/quality choices. Discuss in detail the econometric issues in the estimation of these parameters. Propose and describe in detail a method that provides consistent estimates of $\{F_L, F_H\}$.

(c) Suppose that you incorporate unobserved market heterogeneity in fixed costs: $F_{Lm} = F_L + \eta_m^L$ and $F_{Hm} = F_H + \eta_m^H$, where η_m^L and η_m^H are zero mean random variables, and they are common knowledge to the players. Explain which are the additional econometric issues in the estimation of $\{F_L, F_H\}$ when we have these additional unobservables. Propose and describe in detail a method that provides consistent estimates of $\{F_L, F_H\}$ and the distribution of $\{\eta_m^L, \eta_m^H\}$.

PROBLEM I

Consider an extension of Rust's machine replacement model (Rust, 1987) that incorporates asymmetric information in the market of machines. A firm produces at several independent plants (indexed by i) that operate independently. Each plant has a machine. The cost of operation and maintenance of a machine increases with the age of the machine. Let x_{it}.be the age of the machine at plant i and at period t. There are two types of machines according to their maintenance costs: low and high maintenance costs. When the firm's manager decides to buy a machine, he does not observe its type. However, the manager learns this type just after one year of operation. The maintenance cost is: $c_i \ x_{it} + \varepsilon_{it}(0)$ where $c_i \in \{\theta_L, \theta_H\}$ is a parameter and $\varepsilon_{it}(0)$ is a component of the maintenance cost that is unobserved for the researcher. There is a cost of replacing an old machine by a new one. This replacement cost is: $RC + \varepsilon_{it}(1)$ where RC is a parameter, and $\varepsilon_{it}(1)$ is a component of the maintenance cost that is unobserved for the researcher. The firm has decide when to replace a machine in order to minimize the present value of the sume of maintenance and replacement costs. Suppose that the researcher has a random sample of machines.

PROBLEM J

1. **Demand of a differentiated storable product.** Consider a model of demand of a differentiated storable product as in the paper by Hendel and Nevo (*Econometrica,* 2006). Consumers are forward-looking and can hold inventories of the product. The researcher has panel data on several thousand consumers with weekly information on product prices and consumer purchasing decisions over more than one hundred weeks.

    (a) Explain how consumer purchasing behavior depends on consumer inventories and on expectations about future prices. Based on these predictions of the dynamic model, propose a simple test of consumer forward-looking behavior using panel data of prices and consumer purchasing decisions. Discuss the assumptions under which your testing procedure is a valid test of consumer forward-looking behavior.

    (b) Suppose that a researcher ignores consumer forward-looking behavior and estimates a static model of demand of a differentiated product. Explain the implications on estimated elasticities and firm market power of ignoring (when present) consumer forward-looking and stockpiling behavior. Discuss how these biases depend on the stochastic process followed by prices. In particular, discuss the biases when prices follow a Hi-Lo pricing pattern.

    (c) Suppose that we assume that there is not product differentiation in consumption and in inventory holding costs of the storable product such that all the product differentiation occurs at the moment of purchase. Write a consumer utility function with this property. Discuss the implications of this assumption on the structure of the model and on its estimation. Explain in detail how to estimate the parameters in the component of the utility function that incorporates product differentiation and affects brand choice.

    (d) Explain the 'curse of dimensionality' associated with the estimation of this dynamic demand model. Describe the "inclusive values approach" and how it deals with the high dimensionality problem.

    (e) As described above, the researcher observes the histories of consumer purchasing decisions but she does not have data on consumer inventories or on consumption. Describe an approach (either a set of model assumptions or an econometric technique) to deal with this issue in the estimation of the demand of a storable product.

PROBLEM K

2. **Static Game of Oligopoly Competition with Incomplete Information.** *Business Inn* (brand B) and *Quantity Inn* (brand Q) are two 'economy' hotel brands operated by two separate hotel chains. These brands compete with each other in M local lodging markets. Their entry decisions are represented by the binary variables $a_{Bm} \in \{0,1\}$ and $a_{Qm} \in \{0,1\}$, where $a_{im} = 1$ represents entry in market m by brand $i \in \{B, Q\}$. Each market m is characterized by the number of travelers (N_m) and by whether there is a *Business Executives Suite* hotel (brand BES) operating in the market. Brand BES is a luxury hotel brand owned by the same chain that operates *Business Inn*. Let $BES_m \in \{0,1\}$ denote a dummy variable whose value is 1 if market m has a BES. Since a BES is a luxury brand, the presence of this brand in the market does not affect the demand of the two 'economy' brands. However, it does affect the operating cost of *Business Inn* because the 'economy' and the luxury hotel brands can share some of their inputs (e.g., laundry facility) and exploit some economies of scope/scale. The profit of operating the economy brands in market m are written by

$$\pi_{Bm} = \begin{cases} \alpha_{B0} + \alpha_{B1} \ln N_m + \alpha_{B2} BES_m + \delta_B a_{Qm} + \epsilon_{Bm}(1) & \text{if } a_{Bm} = 1 \\ \epsilon_{Bm}(0) & \text{if } a_{Bm} = 0 \end{cases}$$

$$\pi_{Qm} = \begin{cases} \alpha_{Q0} + \alpha_{Q1} \ln N_m + \delta_Q a_{Bm} + \epsilon_{Qm}(1) & \text{if } a_{Qm} = 1 \\ \epsilon_{Qm}(0) & \text{if } a_{Qm} = 0 \end{cases}$$

where $\alpha' s$ and $\delta' s$ are parameters, and $\epsilon_{Bm}(\cdot)$ and $\epsilon_{Qm}(\cdot)$ are i.i.d. over markets and brands with a Type I extreme value distribution. We assume that $\epsilon_{im}(\cdot)$ is private information of firm i. The researcher observes $\{N_m,\ BES_m,\ a_{Bm},\ a_{Qm}\}_{m=1}^{M}$.

(a) Assume that BES_m is exogenous to the entry decisions of the two economy brands. Explain why this assumption can be reasonable despite the fact that a *Business Executives Suite* shares the same input with a *Business Inn*.

(b) Describe a Bayesian Nash Equilibrium (BNE) in this model.

(c) Explain how to estimate the structural parameters of this model by the nested fixed point algorithm. Also, explain why it is difficult to implement this algorithm when the model has more than one equilibria.

(d) Explain how to estimate the same parameters by using a two-step method. What assumption(s) do we need to impose to obtain consistent estimates when the model has more than one equilibria? Explain why this estimator does not work without imposing such assumption(s).

(e) Does the identification of this model rely on the functional form assumptions on the specification of firms' profit functions? Explain.

PROBLEM L

3. **Production function estimation.** Consider a Cobb-Douglas production function (in logs)

$$y_{it} = \beta_l \ l_{it} + \beta_k \ k_{it} + w_{it}$$

where i is the firm index, t is the year index, y_{it} represents log of deflated sales, l_{it} is the log of employment, k_{it} is the log of capital at the beginning of year t, and w_{it} represents total factor productivity. For simplicity, assume that total factor productivity has the following error components structure, $w_{it} = \eta_i + \delta_t + w_{it}^*$ where the component w_{it}^* follows an AR(1) process with parameter $\rho \in (0, 1)$. The researcher has an unbalanced panel of firms with information on y_{it}, l_{it}, k_{it}, and on capital investment, x_{it}. We are interested in the estimation of the parameters β_l and β_k and of the distribution of total factor productivity.

(a) Explain the following two econometric issues in the estimation of production functions: simultaneity and endogenous exit.

(b) Explain how to estimate the parameters β_l and β_k using OLS with time fixed effects. Discuss why this estimator can be biased/inconsistent.

(c) Describe a "dynamic panel data" method for the estimation of the parameters β_l and β_k. Explain the identifying assumptions of this econometric method.

(d) Describe Olley and Pakes method for the estimation of the parameters β_l and β_k. Explain the identifying assumptions of this econometric method.

(e) Consider the same conditions as in the Olley-Pakes model in Question (d) with the only difference that now firms have different prices of capital. For instance, the interest rate that firms pay may depend on the firm's likelihood of failure. Explain why under these conditions, the standard Olley-Pakes method that you described in Question (d) does not deliver consistent estimates for β_l and β_k. Suppose that you have access to data on the credit ratings of these firms, and that those credit ratings are related to the firm's price of capital. Propose an extension of Olley-Pakes method that provides consistent estimates in this context. Explain the new assumptions.

PROBLEM M

1. Consider a supermarket industry and suppose that there are only two firms in the industry, supermarket chains 1 and 2. Competition takes place at the level of local markets. There are M local markets. Consider the following adaptation to this industry of the simultaneous equations model in *Olley and Pakes (Econometrica, 1996).*

$$\begin{array}{ll} \text{Production Function:} & y_{imt} = \alpha_{Li}\ l_{imt} + \alpha_{Ki}\ k_{imt} + \omega_{imt} + e_{imt} \\ \text{Investment Function:} & i_{imt} = f_i\left(k_{imt}, \omega_{imt}\right) \end{array}$$

where $i \in \{1, 2\}$ is the firm subindex, m is the local-market subindex, and t is the time subindex. The rest of the variables and parameters have the same interpretation as in Olley-Pakes. We have a panel dataset for M markets, T periods, and two firms, where T is small and M is large. Following these authors, we assume that labor is a perfectly flexible input and that new investment is not productivity until next period (i.e., time-to-build). We are interested in the estimation of the production function parameters $\{\alpha_{L1}, \alpha_{K1}, \alpha_{L2}, \alpha_{K2}\}$.

   (a) Describe Olley-Pakes method. Explain the role of different assumptions for the consistency of the estimator.

   (b) In the standard application of Olley-Pakes method, most of the sample variation comes from variability across firms, and the asymptotics of the estimator is in the number of firms N. However, in this application there are only two firms, most of the sample variation is across markets, and the asymptotics of the estimators should be in the number of markets M. Explain how these differences with respect to the standard Olley-Pakes framework affect the estimation of the model.

   (c) In Olley-Pakes framework the investment function is $f_t\left(k_{it}, \omega_{it}\right)$, i.e., it varies over time but not over firms. The subindex t in the investment function is implicitly taking into account that the capital stock and productivity of other firms in the industry (i.e., market structure) enter in this investment function. In contrast, in our specification above, the investment function f_i does not vary over time or over markets. Implicitly, we are assuming monopolistic competition. Discuss the role of this assumption. Suppose that given our type of data, we had an investment function f_m that varies over markets. Show that the standard application of Olley-Pakes method to these data and model would provide inconsistent estimates of the parameters of interest. Explain why.

   (d) Consider an investment function $i_{imt} = f_i\left(k_{1mt}, k_{2mt}, \omega_{1mt}, \omega_{2mt}\right)$. Assume that f_1 and f_2 are invertible such that there are inverse functions $\omega_{1mt} = h_1(k_{1mt}, k_{2mt}, i_{1mt}, i_{2mt})$ and $\omega_{2mt} = h_1(k_{1mt}, k_{2mt}, i_{1mt}, i_{2mt})$. Explain how Olley-Pakes approach can be generalized to estimate consistently the parameters of this model.

PROBLEM N

2. **A dynamic game of market entry-exit**

Consider an oligopoly industry characterized by local competition. A researcher has panel data of M local markets over T years, where M is large and T is small. Markets are indexed by m and years are indexed by t. For every market and year, the dataset includes information on: market size, s_{mt}; the number of active firms, n_{mt}; the number of new entrants during the year, en_{mt}; and the number of exiting firms, ex_{mt}. Consider the following model of oligopoly competition in a local market. There are N firms that may operate in the market. A firm in this market can be either active or inactive. The profit of an inactive firm is zero. The profit of an active firm in a market with n competitors is:

$$\Pi_{mt}(n) = s_{mt}\left(\theta_0^{VP} - \theta_1^{VP} n\right) - \theta^{FC} - \varepsilon_{imt} - (1 - a_{imt-1})\theta^{EC}$$

θ_0^{VP}, θ_1^{VP}, θ^{FC}, and θ^{EC} are parameters. a_{imt-1} is the binary indicator of the event "firm i was an incumbent at period $t-1$". ε_{imt} is a component of the fixed operating cost that varies over time, across markets, and across firms, and it is private information of firm i. We assume that ε_{imt} is iid over time, markets, and firms, with a $N(0, \sigma_\varepsilon^2)$ distribution. Market size evolves exogenously over time according to a Markov process with transition probability function $f_s(s_{mt+1}|s_{mt})$. Every period t, firms observe market size, the number of active firms in the market at previous period, and their own private fixed cost, and then they decide simultaneously whether to be active in the market or not. Firms are forward-looking and play strategies that depend only on payoff-relevant state variables. The equilibrium in this model is a Markov Perfect Equilibrium (MPE). Given that firms are identical, up to their private information ε_{imt}, we consider only symmetric MPE.

(a) Describe in detail the structure of a MPE in this model. Derive and explain the following objects in this model: (a.1) the vector of payoff relevant state variables; (a.2) the expected one-period profit; (a.3) the transition probability of the state variables; (a.4) the dynamic decision problem of an incumbent firm and his best response function; (a.5) the dynamic decision problem of a potential entrant and his best response function; (a.6) the best response probability function; (a.7) the MPE as a fixed point of a mapping in the space of firm's choice probabilities.
(b) Let x_{mt} be the vector (s_{mt}, n_{mt-1}). Let $P_0(x_{mt})$ be the probability that a potential entrant chooses to enter in the market, and let $P_1(x_{mt})$ be the probability that an incumbent firm decides to stay in the market. Let $P_{en}(en_{mt}|x_{mt})$ and $P_{ex}(en_{mt}|x_{mt})$ be the probability distributions for the number of entrants and the number of exits conditional on x_{mt}, respectively. Write the distribution $P_{en}(.|x_{mt})$ in terms of the probability $P_0(x_{mt})$, and the distribution $P_{ex}(.|x_{mt})$ in terms of the probability $P_1(x_{mt})$.
(c) Consider the conditional log-likelihood function:

$$l(\theta) = \sum_{m=1}^{M} \log \Pr\left(n_{m2},\, n_{m3},\, ...,\, n_{mT} \mid n_{m1},\, s_{m1},\, s_{m2},\, ...,\, s_{mT}\right)$$

where θ is the vector of structural parameters. (c.1) Write this log-likelihood function in terms of the probabilities $P_{en}(en_{mt}|x_{mt})$ and $P_{ex}(ex_{mt}|x_{mt})$. (c.2) Suppose that for every value of θ the model has a unique equilibrium. Describe in detail a method for the estimation of θ in this model. (c.3) In general, there are values of θ for which the model has multiple equilibria. Describe in detail a two-step method for the estimation of θ. Explain how this method can be extended recursively.
(d) Consider the following extension of the original model. The fixed operating cost in market m is $FC_m = \theta^{FC} + \omega_m$, where ω_m is a zero mean random variable that is common knowledge to all the firms. (d.1) Explain why this model can explain the evidence that, conditional on market size s_{mt}, entry is positively correlated (and exit is negatively correlated) with the number of incumbent firms at the beginning of the year. (d.2) Propose a method to estimate this model.

PROBLEM O

3. A dynamic model of electricity production in a competitive market

Consider a dynamic model of electricity production. Dynamics of electricity production arise due to generator start-up costs. Start-up costs occur whenever a generator is turned on after a period of zero production. Generators are modelled as single firms. In each period, the firm observes the price in the market and the interval of the day. The firm can take one of two actions

$$a_{it} = \left\{ \begin{array}{cc} 1 & \text{if operate in } t \\ 0 & \text{if not operate in } t \end{array} \right\}$$

where i indexes the generator and t a fiften minute time period.

If the firm decides to operate, the firm's output will be one of two levels. If the price is greater than the firm's marginal cost then it will produce at maximum capacity. Instead, if the price is below their marginal cost, then the firm will operate at the minimum level.

$$q_{it} = \left\{ \begin{array}{ll} \max & \text{if } P_t \geq c_i \text{ and } a_{it} = 1 \\ \min & \text{if } P_t < c_i \text{ and } a_{it} = 1 \end{array} \right\}$$

where c_i is the firm's constant marginal cost and P_t is the price for electricity in the generator's zone.

Each period when the firm is operating its profits are

$$\pi(P_t, q_{it}, a_{it}) = \left\{ \begin{array}{ll} (P_t - c_i)q_{it} - FC_i + \varepsilon(1) & \text{if } a_{it} = 1 \text{ and } s_{it} = 1 \\ (P_t - c_i)q_{it} - FC_i - START_i + \varepsilon(1) & \text{if } a_{it} = 1 \text{ and } s_{it} = 0 \\ 0 + \varepsilon(0) & \text{if } a_{it} = 0 \end{array} \right\}$$

where FC_i is the fixed cost of operating generator i; $START_i$ is the cost of starting up generator i; $s_{it} = a_{it-1}$ is the operating state of last period; and the shocks $\varepsilon(a_{it})$ are an iid process distributed as type I extreme value.

The start-up parameters of the model are: FC_i, $START_i$, and c_i. Assume that the c_i are known for each generator and can be independently calculated. Therefore, the parameters FC_i and $START_i$ will be estimated.

Assume that prices follow an AR(1) process described by the distribution $F(P_t|P_{t-1},I_{t-1})$, where I_t is an indicator for each 15 minute interval within a day. The law of motion of I_t is: $I_{t+1} = I_t + 1 - 1(I_t = 96) * 96$. Assume that firms are price takers, the marginal cost of each generator is constant and known, there are no transmission costs, and a generator can costlessly adjust output within its operating range.

Suppose that we have data on capacity, operation (i.e., a_{it}), marginal costs, and prices P_t, for all periods t and producers i. Let (P_t, I_t, s_t) be the state variables of the dynamic problem of a firm and $\beta \in (0,1)$ the discount factor.

(a) Write the Bellman equation representing the dynamic problem of a firm
(b) Characterize the optimal policy for this dynamic problem.
(c) Explain how you can simplify and solve this dynamic problem following Rust's model/methodology.
(d) Let θ be the vector of parameters that we want to estimate. Assume we have N firms and T periods of data, where T is large. Write the likelihood function. Please, clearly specify the meaning and definition of the probabilities you use to construct this function.
(e) Explain what elements in the data could allow an econometrician to estimate $START_i$. Explain how can $START_i$ be separately identified from FC_i.