# Advanced Industrial Organization II
# Problem Set 2

DUE: Sunday Feb 20th at noon

Please typeset your assignment in Latex. For the coding part feel free to use your preferred programming language. Please comment your code clearly, but not excessively, and use low-levels commands. The clarity of your code and the overall assignment will be evaluated. Submit your assignment through Canvas in a single zip file. You can work in a group of two and submit a single file per group.

## Exercise 1

Consider a static entry game between $J$ firms operating in $T$ markets. Upon entry, the profit of firm $j$ in market $t$ depends on the number of entrants and is given by:

$$\pi_{jt}(n_{-jt}) = x_{jt}\beta - \phi_j - \delta_j \log(1 + n_{-jt}) + \epsilon_{jt}$$

where $x_{jt}$ are market-firm specific profits shifter, $\phi_j$ is a firm specific fixed cost, $n_{-jt}$ is the number of competitors in market $t$ and $\epsilon_{jt} \sim F$ is an idiosyncratic shock. Given $n_{-jt}$ firm $j$ enters if and only if $\pi_{jt} \geq 0$.

1. For a given market $t$, find the pure strategy Nash Equilibrium(a) as a function of the realization of $(\epsilon_{1t}, ..., \epsilon_{Jt})$. For simplicity assume that $J = 2$.

2. Suppose you have data on the entry decisions of both firms and some relevant characteristics $\{y_{1t}, y_{2t}, x_{1t}, x_{2t}\}_{t=1}^{T}$. Where $y_{jt} = 1$ if firm $j$ entered market $t$. Let $\prod_{t=1}^{T} l(y_{1t}, y_{2t}|x_{1t}, x_{2t}, \beta, \delta_1, \delta_2)$ the likelihood of the data where, for each $t$, $(y_{1t}, y_{2t})$ is a Nash Equilibrium. Is the likelihood well-defined? Explain.

3. Now suppose that entry is sequential and assume that firm 1 moves before firm 2. Let $\prod_{t=1}^{T} l(y_{1t}, y_{2t}|x_{1t}, x_{2t}, \beta, \delta_1, \delta_2)$ the likelihood of the data where,

for each $t$, $(y_{1t}, y_{2t})$ is a Subgame Perfect Nash Equilibrium. Is the likelihood well-defined? Explain.

4. Assume we have market level data on $T$ markets. For each market $t$, we only observe the number of firms $n_t$ and some market level characteristic $x_t$. Moreover, assume that firms are homogeneous and play a simultaneous entry game in each market with the same entry profits as before:

$$\pi_t(n_t) = x_t\beta - \phi - \delta\log(n_t) + \epsilon_t. \tag{1}$$

   - What are the conditions under which $n_t$ is a Nash Equilibrium?
   - Assume that $\epsilon_t \sim i.i.dN(0,1)$, using the data in pset2_ex1.csv estimate the parameters $(\beta, \phi, \delta)$ using MLE.

# Exercise 2

Following Berry (1992) suppose we have $T$ markets and $K$ heterogeneous potential entrants in each market.[1] Entry profits for potential entrant $i$ in market $t$ are:

$$\pi_{it}(n_t) = \underbrace{\gamma + x_t\beta - \delta\log(n_t)}_{v(x_t,n_t)} + \phi_{it} \tag{2}$$

where $n_t$ is the number of firms that enter market $t$, $v(x_t, n_t)$ is common to all firms in market $t$ while $\phi_{it}$ is the heterogeneous component of profits. We will model firm heterogeneity as follows:

$$\phi_{it} = z_{it}\alpha + \rho\eta_t + \sqrt{1-\rho^2}\epsilon_{it} \tag{3}$$

where $z_{it}$ is an observed firm-market specific characteristic, $\eta_t \sim i.i.dN(0,1)$ is a market specific shock and $\epsilon_{it} \sim i.i.dN(0,1)$ is a firm-market idiosyncratic shock.

Let $T$ be the number of markets, $K = 30$ the number of potential entrants and $(\alpha, \beta, \delta, \gamma, \rho) = (1, 2, 6, 3, 0.8)$ the set of true parameters. Also let $x_t \sim i.i.d\exp(N(0,1))$, $z_{it} \sim i.i.dN(0,2)$, in what follows we will perform a Monte Carlo exercise to recover $\theta$.

- To guarantee equilibrium uniqueness, assume that the entry decision is sequential and that firms enter in order of their heterogeneous profitability. That is,

[1] The potential entrants are the same in each market. For instance, think about $K$ airlines companies that could potential enter $T$ city-pair markets.

for each market $t$, we sort the $K$ potential entrant based on their realized $\phi_{it}$ and allow them to enter sequentially in that order. For a generic market $t$, find the SPNE equilibrium $n_t^*$.

- Under the above assumptions, choose a large $T$ and simulate the model to generate a dataset that for each $t$ contains $x_t$, $z_{it}$, the equilibrium number of entrant $n_t^*$ and an indicator for the identity of each entrant $I_{it} = \mathbf{1}\{i \text{ enters } t\}$.

In what follows, given our simulated data, we will exploit some of the moments conditions implied by our entry model to see if we can recover $\theta = (\alpha, \beta, \delta, \gamma, \rho)$.[2] Note that under our modeling assumptions, for each market $t$ we can decompose the observed number of entrants and their identities as follows:

$$\begin{aligned}
n_t^* &= \mathbb{E}\left[n_t^*(\theta, x_t, z_t)|x_t, z_t, \theta\right] + u_t \\
I_{it} &= \Pr\{I_{it}(\theta, x_t, z_t) = 1|x_t, z_t, \theta\} + \xi_{it} \quad i = 1, ..., K
\end{aligned}$$

where $z_t = (z_{kt})_{k=1}^K$ and the expectations are taken over the distribution of $(\eta_t, \epsilon_{it})$.

- Under the assumption that our model is correctly specified, the above implies the following conditional moments restrictions:

$$\mathbb{E}[u_t|x_t, z_t] = 0 \quad \text{and} \quad \mathbb{E}[\xi_{it}|x_t, z_t] = 0 \text{ for } i = 1, ..., K.$$

  Write down some moments that you want to use to estimate $\theta$.

- Code up the GMM objective using your chosen moments. Note that your moments will not have a closed form solution so you need to use simulations. For instance, suppose you want to use, among the others, the following moment $\mathbb{E}[(I_{it} - \Pr\{I_{it}(\theta, x_t, z_t) = 1|x_t, z_t\})z_{it}] = 0$ then you need to take say $S$ draws from the distribution of shocks e.g., $\eta_t^{(s)}$ and $\epsilon_t^{(s)} = (\epsilon_t^{(s)})_{i=1}^K$ for $s = 1, ..., S$ to approximate

$$\Pr\{I_{it}(\theta, x_t, z_t)|x_t, z_t\} \approx \frac{1}{S}\sum_{s=1}^{S} I_{it}(\theta, x_t, z_t, \eta_t^{(s)}, \epsilon_t^{(s)})$$

  where $I_{it}(\theta, x_t, z_t, \eta_t^{(s)}, \epsilon_t^{(s)})$ is computed by finding the SPNE of the model under draw $s$ and checking whether or not firm $i$ enters market $t$.

[2] While tempting because we know the distribution of the errors, MLE is quite unpractical here (why?).

- For each of the parameters in $\theta$ plot your GMM objective over a grid around the true parameter while holding all the other parameters at their true value. For instance, if you plot GMMobjective($\alpha; \gamma = 3, \beta = 2, \delta = 6, \rho = .8$) for $\alpha \in [-5, 5]$ and your GMM objective is correct you should see a local minimum at $\alpha = 1$. The same should be true for the other parameters.

- Repeat your estimation several times and report an histogram of your estimates for each parameter together with the initial condition you used.

- Report your results for different initial conditions.

**Remark (1):** Very likely, your GMM objective won't be differentiable so you need to use some gradient-free optimization.[3]

**Remark (2):** Estimation can be sensitive to your initial condition. You don't want to start too far from the true value, try playing around with that.

# Exercise 3

In the file ps2_ex3.csv you will find mileage data for the bus managed by one Harold Zurcher. For this exercise, there are no covariates. Each time period, Harold Zurcher chooses to perform maintenance on the bus, or to replace the engine. Let his flow utilities be given by the following function

$$u(x_t, d_t) + \epsilon_{a,t} = \begin{cases} -\theta_1 x_t - \theta_2 \left(\frac{x_t}{100}\right)^2 + \epsilon_{0,t} & \text{if } d_t = 0 \\ -\theta_3 + \epsilon_{1,t} \text{ if } d_t = 1 \end{cases} \tag{4}$$

Where $x_t$ is the current mileage of the bus, $d_t$ is the choice of Harold Zurcher, and $\theta$ is a vector of parameters. Each choice also contains unobserved utility $\epsilon_{a,t}$ that are distributed independent T1EV.

Harold Zurcher maximizes his lifetime discounted utility, discounted by $\beta$, where the state $x_t$ evolves according to

$$p(x_{t+1}|x_t, d_t) = \begin{cases} g(x_{t+1} - 0) & \text{if } d_t = 1 \\ g(x_{t+1} - x_t) & \text{if } d_t = 0 \end{cases} \tag{5}$$

That is, replacing the engine regenerates the mileage to 0.

[3] Depending on your programming language, any version of the Nelder-Mead optimization routine should be fine.

1. How can you recover engine replacement from the mileage data? Store these decisions in a separate variable.

2. Discuss the conditional independence assumption.

3. Discretize the domain of $x_t$ into $K$ chunks. Estimate the Markov Transition probability $p(x_{t+1}|x_t, d_t)$. This should be two $K \times K$ stochastic matrices, depending on action $d_t$. Make your own choice of $K$.

4. Define the expected value function $EV(x, d) = \int V_\theta(y, \epsilon)p(d\epsilon)p(dy|x, d)$. Show the following fixed point equation holds:

$$EV(x, d) = \int \log\left(\sum_j \exp(u(y, j) + \beta EV(y, j))\right) p(dy|x, d) \tag{6}$$

5. Derive the conditional choice probabilities using $EV(x, d)$ and $\theta$.

6. Reduce the state space of $EV$ using the regenerative property.

7. Rewrite the fixed point equation as a matrix equation.

8. Write a function that solves the fixed-point equation using Rust's algorithm.

9. Write a function that computes the likelihood of the sample for any parameter $\theta$.

10. Estimate the model parameters $\theta$. Use $\beta = .999$