Assignment # 2

Ian Jaeyeon Kim

1 Simulation Part

Question 1

- **Analytic solution of $\pi_i(N_t, x_t)$**

$$\begin{aligned}
\pi_i(q_i; N_t, x_t) &= p \cdot q_i - fc \\
&= [(10 + x_t) - \sum_{j=1}^{N_t} q_j] \cdot q_i - fc \\
&= (10 + x_t) \cdot q_i - q_i^2 - (\sum_{j \neq i} q_j) \cdot q_i - fc \\
\Rightarrow \frac{\partial \pi_i(q_i; N_t, x_t)}{\partial q_i} &= (10 + x_t) - 2q_i - (\sum_{j \neq i} q_j) \overset{!}{=} 0 \Big|_{q_i = q_i^*}
\end{aligned}$$

Firms are homogeneous and facing the same best response curve $\therefore q_j^* = q_{NE}^* \, \forall j$

$$\begin{aligned}
&\Rightarrow (10 + x_t) - 2q_{NE}^* - (N_t - 1) \cdot q_{NE}^* = 0 \\
&\Rightarrow q_{NE}^*(N_t, x_t) = \frac{10 + x_t}{N_t + 1} \\
\Rightarrow \pi_i^*(N_t, x_t) &= (10 + x_t - N_t \frac{10 + x_t}{N_t + 1}) \frac{10 + x_t}{N_t + 1} - fc \\
&= (\frac{10 + x_t}{N_t + 1})^2 - fc
\end{aligned}$$

- **One minor comment about $\pi_i(N_t, x_t)$ matrix: option of 'idling'**

Table below shows the associated profit for each pair of (N_t, x_t). Due to the per period fixed cost, there are some negative values in the matrix. I wanted to mention one thing about those

$N_t \backslash x_t$	**-5**	**0**	**5**
1	1.25	20	51.25
2	-2.222	6.111	20
3	-3.437	1.25	9.0625
4	-4	-1	4
5	-4.3055	-2.2222	1.25

negative values. It is natural for an incumbent to choose 'not producing' option, when the firm is expected to have a negative per period profit if it produces. However, if we allow that 'idling' option for an incumbent, first it will add another strategic option for an incumbent. And plus, we can not distinguish the following two situations with the 'idling' option: (1) Three firms are in the market and all of them produce, (2) Four firms are in the market, three of them produce after observing the other one firm decides not to produce. This will make the remaining part not tractable, because we will focus on the number of firms in the market as the equilibrium outcome. So throughout this assignment, I will not allow 'idling' option for incumbents.

Question 2 Note that this dynamic entry/exit game is a bayesian game, because players draw their types μ_i and γ_e privately. And for the associated bayesian nash equilibrium strategy, we will assume cutoff rules for the following **intuitive** reasons:

- **Incumbent's exit decision**
  Let's call an incumbent's exit choice as 'outside option' and remaining choice as 'inside option'. Then observe the followings:

  1. Outside option's value, which is the firm's scrap value μ_i, is increasing in μ_i.
  2. Inside option's value is flat with respect to μ_i. That is because the value of inside option is determined by other model primitives($\pi_i(\cdot,\cdot)$, transition matrices, β), which do not depend on the iid draw of μ_i.

  Hence, an incumbent will exit the market when the outside option's value exceeds the inside option's value. Therefore, incumbents will follow a cutoff rule.

- **Entrant's entry decision**
  Let's call an entrant's not entering choice as 'outside option' and entry choice as 'inside option'. Now observe the followings:

  1. Outside option's value, which is fixed as 0, is constant with respect to γ_e
  2. Inside option's value is decreasing in γ_e. That is because, if we denote an entrant's expected value from playing in the market from tomorrow with $\overline{V}$(which is orthogonal to iid draw γ_e), inside option's value becomes $\overline{V} - \gamma_e$.

  Hence, whenever the value from inside option($\overline{V} - \gamma_e$) is lower than 0, an entrant will not enter the market. Therefore, entrants will follow a cutoff rule.

Question 3

- **What, conceptually, $\bar{V}(N_t, x_t)$ measures?**
  Let's start from $V(N_t, x_t, \mu_{it})$. $V(N_t, x_t, \mu_{it})$ is the expected value of an incumbent who has drawn its μ_{it} as its scrap value under (N_t, x_t). So this is an **ex-post** value in the sense that

it is calculated after μ_{it} is drawn. Now, let's integrate out μ_{it} to come up with $\bar{V}(N_t, x_t)$. Note that $\bar{V}(N_t, x_t) = \mathbb{E}_{\mu_{it}}[V(N_t, x_t, \mu_{it})]$. Hence, $\bar{V}(N_t, x_t)$ means the expected value of an incumbent, who has not drawn its scrap value yet under (N_t, x_t). Conceptually, it measures the **ex-ante** maximum value, an incumbent can expect, when the number of firms in the market is N_t(including me), and the demand shock is x_t.

- **How does it relate to the alternative-specific value functions of Rust(1987)?**
  This ex-ante value of being an incumbent, conditional on the pair of state (N_t, x_t), enables us to write down the **alternative specific value functions** of incumbents and entrants as follows:

$$V^{inc}(N, x, \mu) = \max\{\underbrace{\mu}_{\text{exit}}, \underbrace{\pi(N, x) + 0.9\mathbb{E}_{N',x'}[\bar{V}(N', x')|N, x, \text{remain}]}_{\text{remain}}\}$$
$$V^{ent}(N, x, \gamma) = \max\{\underbrace{0}_{\text{stay out}}, \underbrace{-\gamma + 0.9\mathbb{E}_{N',x'}[\bar{V}(N', x')|N, x, \text{enter}]}_{\text{enter}}\}$$

  Unlike the Rust(1987), where the author attached choice specific cost shocks following multivariate extreme values to reduce Zurcher's problem to a dynamic logit, in this assignment we will directly use the normality of μ and γ and the cutoff rule strategy to bring the model to the data.

Question 4 This question is the procedure of formally writing down the alternative specific value functions of incumbents and entrants.

$$V(N_t, x_t, \mu_{it}) = \max_{d_{it} \in \{0,1\}} (1 - d_{it})\mu_{it} + d_{it}\Big[\pi(N_t, x_t) + 0.9\mathbb{E}_{N_{t+1},x_{t+1}}[\bar{V}(N_{t+1}, x_{t+1})|N_t, x_t, d_{it} = 1]\Big]$$
$$V^E(N_t, x_t, \gamma_{it}) = \max_{e_{it} \in \{0,1\}} e_{it}\Big[-\gamma_{it} + 0.9\mathbb{E}_{N_{t+1},x_{t+1}}[\bar{V}(N_{t+1}, x_{t+1})|N_t, x_t, e_{it} = 1]\Big]$$

Question 5

- **STEP(a)** Note that for each $x_t \in \{-5, 0, 5\}$, $\overline{\gamma}(N_t, x_t)$ is defined on $N_t \in \{0, 1, 2, 3, 4\}$, whereas $\overline{\mu}(N_t, x_t)$ and $\overline{V}(N_t, x_t)$ are defined on $N_t \in \{1, 2, 3, 4, 5\}$.
  - $\overline{\gamma}(N_t, x_t)$: When $N_t = 5$, a potential entrant is not allowed to enter. Hence, $\overline{\gamma}(5, x_t)$ is not defined.
  - $\overline{\mu}(N_t, x_t)$: When $N_t = 0$, not a single firm exists in the market. Therefore, no one will draw its scrap value. Hence, $\overline{\mu}(0, x_t)$ is not defined.
  - $\overline{V}(N_t, x_t)$: Throughout the procedure for both incumbents and a potential entrant, $\overline{V}$ is used to calculate expected continuation payoff **conditional on** remaining/entry decision. Hence $\overline{V}$ will be calculated conditional on the presence of at least one firm in the market. Therefore, $\overline{V}(N_t, x_t)$ is defined on $N_t \in \{1, 2, 3, 4, 5\}$

Hence, at first, I will make $\overline{\gamma}(N_t, x_t), \overline{\mu}(N_t, x_t)$ and $\overline{V}(N_t, x_t)$ as 5×3 matrices where 1 is assigned as its initial value for individual elements of the matrices.

- **STEP(b)** For each $x_t \in \{-5, 0, 5\}$, we need the transition matrices $\mathbb{P}(N_{t+1}|N_t, x_t, d_{it} = 1)$ (let's say M1) and $\mathbb{P}(N_{t+1}|N_t, x_t, e_{it} = 1)$ (say M2). The difference between M1 and M2 deserves some discussions. M1 and M2 are different with respect to its shape and elements:

  1. Difference in the shape of matrices:
     As we noticed from the STEP(a), M1 transfers $N_t \in \{1, 2, 3, 4, 5\}$ to $N_{t+1} \in \{1, 2, 3, 4, 5\}$. Whereas M2 transfers $N_t \in \{0, 1, 2, 3, 4\}$ to $N_{t+1} \in \{1, 2, 3, 4, 5\}$.
  2. Difference in elements of matrices comes from the difference in where each matrix conditions on (This part is further clarified in the section 1 of appendix):
     M1 is the transition matrix of N, in the perspective of an incumbent firm i who decided to remain in the market. Therefore, to predict N_{t+1}, firm i has to consider both $(N_t - 1)$ incumbents and the potential entrant's decisions. So M1 will be a function of $(N_t - 1)$ number of μ draws, γ draw, and the cut off values $\bar{\mu}(N_t, x_t), \bar{\gamma}(N_t, x_t)$.
     In contrast, **M2** is the transition matrix of N, in the perspective of an entrant e who decided to enter the market. Hence, in predicting N_{t+1}, e doesn't have to consider the potential entrant's decision, because it is already fixed as 1 by itself. Therefore, unlike M1, M2 will be simply a function of N_t number of μ draws and the cutoff value $\bar{\mu}(N_t, x_t)$.

  Remaining detailed steps for part (b) are further clarified in the appendix.

- **STEP(c)** This part is straightforward. This part is further explained in `simulation.py`

- **STEP(d)** LHS of the suggested updating rules are cutoff values and RHS stand for the expected value of inside options(remaining in the market/entering the market) for incumbents and entrants. At the equilibrium level of cutoff values, both incumbents and entrants are indifferent between outside and inside options. Which means the equilibrium level of cutoff values should exactly compensate the expected values from remaining/entering the market for incumbents and entrants respectively. Hence, the indifferent condition between inside and outside options at the equilibrium cutoff values could be our updating rules.

- **STEP(e)**

$$\begin{aligned}
\bar{V}'(N_t,x_t) &= \int_{\mu_{it}} V(N_t,x_t,\mu_{it})dP(\mu_{it}) \\
&= \mathbb{E}_{\mu_{it}}[V(N_t,x_t,\mu_{it})] \\
&= \mathbb{P}(\mu_{it} \geq \bar{\mu}') \cdot \mathbb{E}_{\mu_{it}}[V(N_t,x_t,\mu_{it})|\mu_{it} \geq \bar{\mu}'] + \mathbb{P}(\mu_{it} < \bar{\mu}') \cdot \mathbb{E}_{\mu_{it}}[V(N_t,x_t,\mu_{it})|\mu_{it} < \bar{\mu}'] \\
&= [1-\Phi(\frac{\bar{\mu}'-\mu}{\sigma_\mu})] \cdot \mathbb{E}_{\mu_{it}}[\mu_{it}|\mu_{it} \geq \bar{\mu}'] + \Phi(\frac{\bar{\mu}'-\mu}{\sigma_\mu}) \cdot \mathbb{E}_{\mu_{it}}[\bar{\mu}'|\mu_{it} < \bar{\mu}'] \\
&= [1-\Phi(\frac{\bar{\mu}'-\mu}{\sigma_\mu})] \cdot [\mu + \sigma_\mu(\frac{\phi(\frac{\bar{\mu}'-\mu}{\sigma_\mu})}{1-\Phi(\frac{\bar{\mu}'-\mu}{\sigma_\mu})})] + \Phi(\frac{\bar{\mu}'-\mu}{\sigma_\mu}) \cdot \bar{\mu}'
\end{aligned}$$

  After replacing $\bar{\mu}'$ with $\pi(N_t,x_t)+\Psi_1(N_t,x_t)$, we get the required updating rule for $\bar{V}'(N_t,x_t)$. Thanks to the normal assumption, we can easily integrate out the private shock.

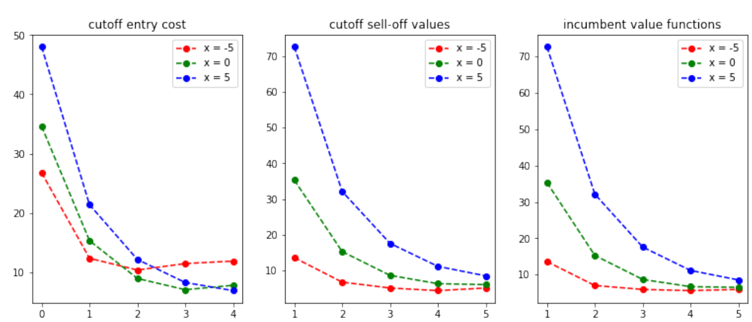
- **STEP(f)** This step is straightforward. It is explained in `simulation.py`

- **STEP(g)** In the updating procedure, I applied the concept of learning ratio with the size of 0.4 to make the updating procedure conservative. (Actually, without the learning rate, the iteration didn't converge).

$$\begin{aligned}
\bar{\mu}'(N_t,x_t) &= 0.6\Big(\bar{\mu}(N_t,x_t)\Big) + 0.4\Big(\pi(N_t,x_t) + \Psi_1(N_t,x_t)\Big) \\
\bar{\gamma}'(N_t,x_t) &= 0.6\Big(\bar{\gamma}(N_t,x_t)\Big) + 0.4\Big(\Psi_2(N_t,x_t)\Big)
\end{aligned}$$

Question 6 As an experiment, I initialized the cutoff matrix and associated value matrix with five different values (I filled the matrices with 1,2,3,4 and 5 for each trial). But the resulting converged matrices didn't change across different initial points. In other words, I couldn't find evidence of multiple equilibria. This is because, in this practice, we are focusing on the total number of firms in the market rather than who's in the market as suggested by Berry(1992).

Question 7 Graph below presents some comparative statics.



1. Cutoff entry cost: First, we can observe that under a given x_t, as N_t goes up, cutoff entry cost goes down. That is because, entrants can endure a competitive markets only when its

entry cost is low. And under a fixed N_t, cutoff entry cost goes up as x_t increases, which is also intuitive. Entrant can expect more value from a market with a higher demand, so under a given level of competition, entrant can endure larger entry cost with higher demand shock.

2. Cutoff sell-off value: As the entrant, incumbents will more likely to remain in the market when the market condition is more favorable (low competition, high market demand). And this is well described in the graph above. Under a fixed level of x_t, cutoff sell-off value goes down as N_t increases. Also, under a fixed number of N_t, cutoff sell-off value goes up as x_t increases.

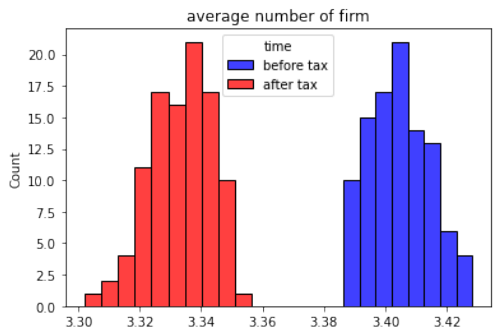
3. Incumbent value function: Explanation for this part is same with the above. Incumbent can expect more value from staying in the market when the market condition is more favorable. And this pattern is described in the graph above.

Question 8

- **Forward Simulation**
  1. x_t process: x_t is driven exogenously by the markov transition matrix.
  2. N_t process: N_t is endogenized. Optimal decision of firms comes in this part.
     - When $N_t = 0$, potential entrant is the single decision maker.
       (a) Draw γ_{it} from $\mathcal{N}(\mu_\gamma, \sigma_\gamma)$.
       (b) $N_{t+1} = N_t + \mathbb{1}(\gamma_{it} \leq \overline{\gamma}(N_t, x_t))$
     - When $N_t \in \{1, 2, 3, 4\}$, both incumbents and the potential entrant make their decisions.
       (a) Draw N_t many μ from $\mathcal{N}(\mu_\mu, \sigma_\mu)$.
       (b) Draw γ_{it} from $\mathcal{N}(\mu_\gamma, \sigma_\gamma)$.
       (c) $N_{t+1} = N_t -$ # of exit decisions $+ \mathbb{1}(\gamma_{it} \leq \overline{\gamma}(N_t, x_t))$
     - When $N_t = 5$, only the 5 incumbents make their decisions.
       (a) Draw 5 many μ from $\mathcal{N}(\mu_\mu, \sigma_\mu)$.
       (b) $N_{t+1} = N_t -$ # of exit decisions
- **What is the average number of firms in the market across 10,000 periods?**
  To be more precise, under a fixed path of x_t, I did 100 times of MC simulations and then take the average of 100 many average number of firms. In other words, I report the number $\bar{N} = \frac{1}{100}\sum_{s=1}^{100}(\frac{1}{10{,}000}\sum_{t=1}^{10{,}000} N_t^s)$ where s stands for each MC simulation. From the simulation, I got $\bar{N} = 3.4043$.

Question 9

- **Result**: Implementation of 5 unit of entry tax will shift down the entrant's expected value from entering the market, which means the tax will lower the equilibrium entry threshold $\overline{\gamma}(N_t, x_t)$. After modifying the cutoff entry cost, as I did in Question 8, I conducted 100 times of MC simulations to get the number $\bar{N}^{tax} = \frac{1}{100}\sum_{s=1}^{100}(\frac{1}{10{,}000}\sum_{t=1}^{10{,}000} N_t^s)$. As a result, I got $\bar{N}^{tax} = 3.3337$ (In doing so, I used the same path of x_t).



  Difference in $\bar{N}$ and $\bar{N}^{tax}$ turns out to be significant. As the graph above shows, the distribution of 100 simulated results for pre and post tax didn't overlap. And the associated t-test(t-stat:49.46) substantiated that the implementation of tax significantly lowers the equilibrium number of firms in the market.

- **Interpretation**: There are two different driving forces for this result. First, when it comes to the **entrant side**, entry tax will lower the probability of entry because the tax implementation lowers the value of entry option by the amount of tax. Hence, the effect of tax on entrants will lower the equilibrium number of firms in the market. However, when it comes to the **incumbent side**, entry tax will elevate the probability of remaining in the market because the lowered threat from the entry will make $\mathbb{P}(N_{t+1}|N_t, x_t, d_{it} = 1)$ more favorable to the incumbent. Therefore, the effect of tax on incumbents will elevate the equilibrium number of firms in the market. Resulting number from the MC simulations above shows that the tax's direct impact on entrant side was bigger than its secondary impact on incumbent side.

- **Coherency under multiple equilibria**: Under multiple equilibria, we can not proceed to counterfactual exercise. That is because, under multiple equilibria, mapping from parameters to outcomes is not unique.

2 Estimation Part

Question 10 To estimate the probability of remaining/entry for each state (N_t, x_t), I used non-parametric frequency estimator. Small size of bins and plenty number of observations allowed me to use the simple but flexible frequency estimator. I did not use other alternatives, such as probit or logit, to avoid any distributional assumption and make first stage as flexible as possible.

Question 11

- From Question 10, we estimated remaining/entry probability from the data. We can also calculate counterpart remaining/entry probability from the model. Let $\Lambda(N_t, x_t; \theta)$ and $\Lambda^E(N_t, x_t; \theta)$ stand for PDV of remaining and entry at (N_t, x_t) under the belief θ. Then, from the cutoff rule and normal assumptions, model predicted remaining and entry probability can be calculated as follows:

$$\text{Model predicted remaining prob under } N_t, x_t = \Phi(\frac{\Lambda(N_t, x_t; \theta) - \mu}{\sigma_\mu})$$
$$\text{Model predicted entry prob under } N_t, x_t = \Phi(\frac{\Lambda^E(N_t, x_t; \theta) - \gamma}{\sigma_\gamma})$$

  Hence, if we can estimate the PDV $\Lambda(N_t, x_t; \theta)$ and $\Lambda^E(N_t, x_t; \theta)$, we can search for true θ using minimum distance approach.

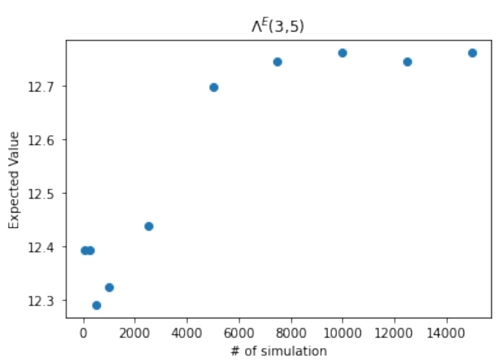
- Part (a), (b): To estimate PDV for a given θ, we will use the forward simulation method following BBL. To do so, we need to recover $\overline{\mu}(N_t, x_t; \theta)$ and $\overline{\gamma}(N_t, x_t; \theta)$ for each search step θ. I tried two different methods to recover the equilibrium cutoff.

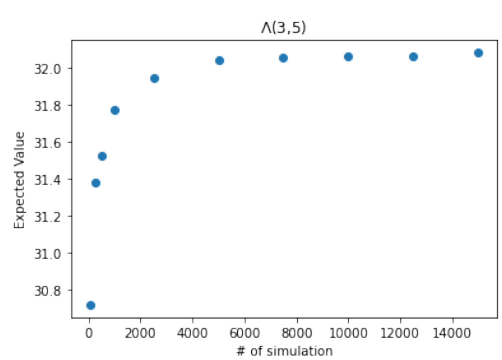
  1. Use result from Question 10 (First method):

$$\overline{\mu}(N_t, x_t; \theta) = \mu + \sigma_\mu \Phi^{-1}(\hat{d}(N_t, x_t))$$
$$\overline{\gamma}(N_t, x_t; \theta) = \gamma + \sigma_\gamma \Phi^{-1}(\hat{e}(N_t, x_t))$$

  2. Solve the model as in Question 5 (Second method)

  First method coincides with the spirit of two step approach, but after all estimation procedure using first method didn't work. I think the first stage estimates was not precise enough. That is why I came up with the second method, which was available in our simple context. Some random experiment results depicted in graphs below showed that PDV value gets stable with more than 5,000 paths, but to decrease the computation time, I simulated 1,000 paths for each search of θ.





- Part (c)

  - Why would this be a reasonable criterion function to use to estimate θ?
    $\Phi(\cdot)$ term stands for the model predicted remaining(entry) ratio under each N_t, x_t, whereas $\hat{d}(N_t, x_t)$ and $\hat{e}(N_t, x_t)$ stand for the estimated remaining and entry probability. Therefore, minimization procedure of $Q_n(\theta)$ fits the model to the data.
  - Why should this object be small at the true parameter value?
    Incumbents and entrants will make their decision under the true belief θ^*. And the data we have is revealed preference of the true parameter. Hence, under the true parameter value, model will spit out the most close value with the data moments.

Question 12 For the two different methods I described above, I used the 'Nelder-Mead' numerical minimization algorithm, because the objective function is not smooth. And I imposed bounds of $(0, 10)$ for each of the parameter to concentrate on positive values and to decrease running time. And for the initial value, I used $(4, 2, 6, 2)$.

Question 13 I double checked my code, but unfortunately as the table below shows, the first method didn't work. The fact that the second method relatively well recovered parameters implies that first stage estimates matters for the second stage estimation.

method\param	γ	σ_γ	μ	σ_μ
First Method	4.07	5.74	0	10
Second Method	3.86	2.00	5.44	2.24

Question 14 I will compare and contrast our estimator to BBL across its first and second step.

1. **First Step**: In the first step, BBL firstly estimates policy function and transition matrix. And using the estimated results, value function is calculated from the forward simulation. Our estimator mostly follows the first step of BBL. One different thing is estimation of transition matrix. Unlike BBL, in our practice, we calculate the transition matrix based on each player's equilibrium perception and normality assumptions for μ and γ.

2. **Second Step**: In the second step, BBL search for θ that rationalizes the primitives from the first stage using the definition of Markov Perfect Equilibrium. To use the inequality condition of MPE, they perturb the policy function. Unlike the BBL, in this practice, we don't use the inequality condition. Instead, we used the minimum distance approach.

Context of our practice enables us to utilize minimum distance approach. Our entry/exit game, where entry is not allowed under $N_t = 5$ and $x_t \in \{-5, 0, 5\}$ is an exogenous markovian process, has only 18 $(= 6 \times 3)$ possible states. With this small set of state couples, 10,000 periods of data points allows us to recover a consistent policy function. Finally, normality assumption on μ and γ

allowed us to calculate the entry/exit probability from the model.
Minimum distance approach is relatively simple compared to the second stage of BBL. We don't have to consider the way to perturb the policy function and the objective function is simpler than the one from the MPE inequality condition. However, our minimum distance approach implicitly assumes point identification. So it is more context specific. Under the circumstances where partial identification is more suitable(where firm heterogeneity matters a lot), we can't apply our minimum distance approach. Furthermore, BBL's approach is more general in the sense that it doesn't require one to come up with moment conditions as we used in this practice.

Question 15 Apart from the slight difference in timing assumption in the model (in POB, scrap value is realized tomorrow, after taking per period profit today), POB and this assignment are quite similar. Most importantly, both practices share the followings:

1. Each equilibrium strategy generates a finite state of (n, x), where $x = z$ in POB
2. The distribution of (n', x') only depends on the current (n, x)
3. Associated data shows recurrent pattern

So from the *proposition 1* in POB, in both practices, existence of a consistent policy function is guaranteed and based on this fact, both practices adopt minimum distance approach that minimizes distance between model predicted exit/entry probability and estimated exit/entry probability from the data. However, there exists difference in calculating the model predicted probability.

1. In our practice, we adopted nested FXP approach in the sense that we calculated equilibrium continuation value(by simulation) for each belief θ given from the outer loop. And then we calculate exit/entry probability using the normality assumptions on μ and γ.
2. In POB, they assumed exponential distribution for scrap value. With this assumption, incumbent's continuation value, $\hat{VC}(\theta)$, is solved in a closed form in a linear fashion. So, with the non parametrically estimated exit probability and transition matrix, they could easily get the incumbent's continuation value for a given θ. Since entrant's entry value nests $\hat{VC}(\theta)$, they did not have to go through the simulation procedure for entrant side. Finally, with the exponential distribution assumption, model predicted exit/entry value is calculated.

The remaining estimation procedure is essentially same. One thing suggested by POB is to aggregate the state specific moment conditions to average out the first stage error. But I omit any further clarification.

Question 16 We need to control the error from the first stage estimates, $\hat{d}(N_t, x_t), \hat{e}(N_t, x_t)$, to measure the pure error in $\hat{\theta}$. To do so, we may follow the following parametric bootstrap procedure: (Intuition: a set of simulated random data is already conditioned on the first stage error, so we may regard the variation in the estimates from the set of simulated data comes from the second stage)

1. Do the entire estimation to get $\hat{d}(N_t, x_t), \hat{e}(N_t, x_t)$ and $\hat{\theta}^0$.
2. Based on $\hat{d}(N_t, x_t), \hat{e}(N_t, x_t)$ and $\hat{\theta}^0$, simulate S many random data.
3. For each generated data, do the estimation procedure to get $\hat{\theta}^s \ \forall s = 1, \cdots, S$
4. Report the sample variance of S many estimates.