Industrial Organization II (ECO 2901)
Winter 2020. Victor Aguirregabiria

Final Exam
Due on Thursday, April 2nd, 2020 [before 11:59pm]

INSTRUCTIONS.
1. Write your answers electronically in a word processor.
2. For the answers that involve coding, include in the document the code that you have used to obtain your empirical results, indicating the language of the code (e.g., R, Matlab, Julia, Python) as well as the version of the software/compiler you have used.
3. Convert the document with your answers and code to PDF format and submit the PDF online via Quercus.
4. You should submit your exam before 11:59pm on Thursday, April 2nd, 2020.
5. The exam should be written individually.

The total number of marks is 100.

This exam deals with **estimation and counterfactual experiments** in a **dynamic network game with incomplete information**. More specifically, the model deals with the diffusion of a virus, confinement decisions, and public policies that may reduce the speed of diffusion. The first part of this document describes the model. The second part presents the Questions you need to answer.

---

1 Model

1.1 Framework

1.1.1 Social network

The society consists of I individuals that we index by $i \in \mathcal{I} = \{1, 2, ..., I\}$. Time is discrete and indexed by $t \in \{0, 1, ...\}$. For simplicity, we consider a society where individuals are homogeneous. Each individual i has a *set of social contacts* that we represent with $\mathcal{C}_i$. The set $\mathcal{C}_i$ contains all the people with whom individual i has *physical contact* in a typical day, e.g., co-workers, friends, neighbors, household members.

The combination of all the sets of social contacts, $\mathbb{C} \equiv \{\mathcal{C}_i : i \in \mathcal{I}\}$, describes the *social (physical) network* in this society. This network is an exogenous primitive in the model. The network can vary across societies because cultural reasons, social norms, transportation infrastructures, geographic features, etc.

1.1.2 Health status

Variable x_{it} describes the state of individual i at period t according to her relationship with the virus. This variable x_{it} can take five possible values, $\mathcal{X} \equiv \{-3, -2, -1, 0, 1\}$. State $x_{it} = 0$ means that the individual has not been infected by the virus. State $x_{it} = -1$ represents that the individual is infected and can transmit the virus to other people but has not developed symptoms yet. This state is not observable to the own individual or to others. State $x_{it} = -2$ means that the individual is infected and has symptoms. We assume that this state is common knowledge to all the individuals. Furthermore, we assume that the health system is such that individuals in this state are forced to be isolated, either in a hospital or at home. State $x_{it} = -3$ represents that the individual has died because the virus infection. Finally, state $x_{it} = 1$ means that the individual is immune and cannot transmit the virus.

1.1.3 Information structure

The assumptions about individuals' information are key for the predictions of the model on individuals' behavior and on the speed of diffusion of the virus.

Let $\widetilde{x}_{it}$ be a measure of health status that does not distinguish between states 0 (healthy) and -1 (infected without symptoms). That is, $\widetilde{x}_{it} = x_{it}$ if $x_{it} \in \{-3, -2, 1\}$ and $\widetilde{x}_{it} = [-1$ or $0]$ if $x_{it} \in \{-1, 0\}$. At period t, an individual i's information set has the following elements.

(i) The value of her own $\widetilde{x}_{it}$.
(ii) The value of $\widetilde{x}_{jt}$ for any other individual in her group of social contacts, $\{\widetilde{x}_{jt} : j \neq i, \, j \in \mathcal{C}_i\}$.
(iii) At the aggregate level, the proportion of individuals in state x at period t. We represent this share as $S_t(x)$, and $\mathbf{S}_t$ is the vector $\{S_t(x) : x \in \mathcal{X}\}$.
(iv) The average probability that an individual in state $\widetilde{x}_{jt} = \{-1 \text{ or } 0\}$ decides confinement at period t. We use Q_t to represent this average probability.

Therefore, the information set of individual i at period t consists of $\{[\widetilde{x}_{jt} : j \in \mathcal{C}_i], \, Q_t, \, \mathbf{S}_t\}$. Below, we extend this information set to include also the individual's private information shocks in the utility function. We represent those shocks as $\varepsilon_{it}(0)$ and $\varepsilon_{it}(1)$.

1.1.4 Individual decision

Every period t, individuals in state $\widetilde{x}_{it} = \{-1 \text{ or } 0\}$ make a binary decision about their social relationships: $a_{it} \in \{0,1\}$. If $a_{it} = 0$, the individual decides to maintain its social relationships as usual having physical contact with people in the set $\mathcal{C}_i$. If $a_{it} = 1$, the individual adopts *confinement*. For simplicity, we assume that confinement means that the individual does not have physical relationship with any other member of the society.

We assume that individuals in state $x_{it} = -2$ (infected with symptoms) are forced to confinement, and individuals in state $x_{it} = 1$ (immune) always choose to have social relationships.

For any value of x, let $n_{it}^{(x)}$ be the number of individuals in set $\mathcal{C}_i$ – other than i – who are in state x and have chosen to socially interact at period t:

$$\left\{ \begin{array}{lrcl} \text{For } x \in \{-1,0\}, & n_{it}^{(x)} & = & \sum\limits_{j \in \mathcal{C}_i, j \neq i} 1\{a_{jt}=0\}\ 1\{x_{jt}=x\} \\ \\ & n_{it}^{(1)} & = & \sum\limits_{j \in \mathcal{C}_i, j \neq i} 1\{x_{jt}=1\} \end{array} \right. \tag{1}$$

1.1.5 Virus infection and mortality risk

We now describe the transition probabilities of health status x_{it}. First, it is important to note that the definition of "a period" in this model corresponds to the *incubation period* of the virus, e.g., 15 days in the case of Covid-19. This has important implications on transition probabilities. In particular, physical contact with other people at period t may generate a new infection at period $t+1$ but, if the individual is not infected at $t+1$, physical contacts at t cannot generate infections at $t+2$ or after.

Let $\mathbf{x}_t$ and $\boldsymbol{a}_t$ be the vectors with the health statuses and confinement decisions, respectively, of all the individuals at period t. Based on the argument in the previous paragraph, we assume that $\mathbf{x}_t$ follows a first order Markov process. We also assume that conditional on $(\mathbf{x}_t, \boldsymbol{a}_t)$ the realization of health statuses at $t+1$ are independent and identically distributed across individuals. That is:

$$\Pr\left(\mathbf{x}_{t+1} | \mathbf{x}_t, \boldsymbol{a}_t, \mathbf{x}_{t-1}, \boldsymbol{a}_{t-1}, ..., \mathbf{x}_0, \boldsymbol{a}_0\right) = \Pr\left(\mathbf{x}_{t+1} | \mathbf{x}_t, \boldsymbol{a}_t\right) = \prod_{i=1}^{I} F\left(x_{i,t+1} | \mathbf{x}_t, \boldsymbol{a}_t\right) \tag{2}$$

We impose the following restrictions on the transition probabilities F.

$$\begin{array}{rlll} \text{(i)} & \text{If } x_{it} = -3, & \text{then} & F(-3|\mathbf{x}_t, \boldsymbol{a}_t) = 1 \\ \text{(ii)} & \text{If } x_{it} = 1, & \text{then} & F(1|\mathbf{x}_t, \boldsymbol{a}_t) = 1 \\ \text{(iii)} & \text{If } x_{it} = -1, & \text{then} & F(-2|\mathbf{x}_t, \boldsymbol{a}_t) = 1 \\ \text{(iv)} & \text{If } x_{it} = 0, & \text{then} & F(-3|\mathbf{x}_t, \boldsymbol{a}_t) = F(-2|\mathbf{x}_t, \boldsymbol{a}_t) = F(1|\mathbf{x}_t, \boldsymbol{a}_t) = 0 \\ \text{(v)} & \text{If } x_{it} = -2, & \text{then} & F(-1|\mathbf{x}_t, \boldsymbol{a}_t) = F(0|\mathbf{x}_t, \boldsymbol{a}_t) = F(-2|\mathbf{x}_t, \boldsymbol{a}_t) = 0 \end{array} \tag{3}$$

Condition (i) says that this society does not have zombies. Condition (ii) establishes that immunity is an absorbing state.[1] Condition (iii) says that an individual infected but without symptoms at period t will develop symptoms at $t+1$ with probability one. This is a simplifying assumption. Condition (iv) says that an individual who is healthy at period t can be only in two possible states at period $t+1$: stay healthy, or get infected without symptoms. Finally, condition (v) means that once an individual has developed symptoms, it has only two possible transitions: recover and become immune, or die.

Under conditions (i) to (v), there are only two free probabilities in the transition function F: (a) the *risk of infection*, $F(-1|x_{it}=0, \mathbf{x}_t, \boldsymbol{a}_t)$; and (b) the *mortality rate*, $F(-3|x_{it} = -2, \mathbf{x}_t, \boldsymbol{a}_t)$. We assume that the *mortality rate* does not depend on the state $(\mathbf{x}_t, \boldsymbol{a}_t)$. We use π^m to represent the mortality rate:

$$F(-3|x_{it} = -2, \mathbf{x}_t, \boldsymbol{a}_t) = \pi^m \tag{4}$$

The *risk of infection* is an increasing function of the number of infected people that individual i interacts with at period t. Remember that we assume that people with symptoms are forced into isolation. Therefore, the only infected individuals with social interactions are those without symptoms. Let π_{it}^v be the probability of infection $F(-1|x_{it}=0, \mathbf{x}_t, \boldsymbol{a}_t)$, and let $\Lambda(.)$ be the logistic function. Then,

$$\pi_{it}^v = \begin{cases} 0 & \text{if } a_{it} = 1 \\ \\ \Lambda\left(\eta_0 + \eta_1 \; \mathbb{E}_{it}\left[n_{it}^{(-1)}\right]\right) & \text{if } a_{it} = 0 \end{cases} \tag{5}$$

where η_0 and η_1 are parameters, and $\mathbb{E}_{it}(.)$ is the expectation operator conditional on the individual's information at t. Parameter η_1 is positive and it captures how contagious the virus is, i.e., the *reproduction rate* of the virus.

1.1.6 Preferences

Since only individuals at state $\widetilde{x}_{it} = \{-1 \text{ or } 0\}$ can make a choice about social confinement, we only need to specify preferences at this state. Every individual has a utility function, $U_{it}(a_{it})$, that depends on the number of individuals with whom she has social contact, and on the risk of her own infection.

I consider the following utility function:

$$\begin{aligned} U_{it}(a_{it}) \quad &= \quad \alpha(a_{it}) + \beta \;\; (1 - a_{it}) \left[n_{it}^{(-1)} + n_{it}^{(0)} + n_{it}^{(1)}\right] \\ &- \quad \delta \; 1\{x_{i,t+1}(a_{it}) = -1\} + \varepsilon_{it}(a_{it}) \end{aligned} \tag{6}$$

[1] In this model, we ignore the possibility of mutations of the virus or of second infections.

$\alpha(0)$, $\alpha(1)$, β, and δ are parameters.

Parameter $\alpha(a)$ represents the utility of an individual, net of the effects of social contacts and contagion risks, when her confinement decision is a. We expect $\alpha(0) > \alpha(1)$ since confinement – regardless social interactions – reduces the freedom and the set of possible actions of an individual. Parameter β captures the marginal utility of individual i from an additional social contact. We expect β to be a positive parameter. Parameter δ measures individual i's disutility from being infected. Variable $x_{i,t+1}(a_{it})$ represents the realization of the individual's health status at $t+1$ given that the choice at period t is a_{it}. This choice can only affect the transition from $x_{it} = 0$ to $x_{it} = -1$: the rest of the transitions are not affected by the confinement decision. This is why the term in the utility is only the indicator $1\{x_{i,t+1}(a_{it}) = -1\}$. Health status at $t+1$ is unknown at period t. Therefore, individuals form expectations about variable $1\{x_{i,t+1}(a_{it}) = -1\}$ and maximize expected utility.

Variables $\varepsilon_{it}(0)$ and $\varepsilon_{it}(1)$ are private information of individual i. We assume that $\varepsilon_{it}(0) - \varepsilon_{it}(1)$ is independently and identically distributed across individuals and over time with a logistic distribution.

1.2 Equilibrium

An equilibrium in this model is a sequence of probabilities $\{Q_t, \mathbf{S}_t : t \geq 0\}$ such that they are consistent with individuals' beliefs and best responses, and with the transition probabilities π^m and π^v_{it}.

1.2.1 Expected utility and best response probability function

The difference between the expected utilities from confinement and no-confinement is:

$$\widetilde{\alpha} - \beta \left(\mathbb{E}_{it} \left[n_{it}^{(-1)} + n_{it}^{(0)} \right] + n_{it}^{(1)} \right) - \delta \, \mathbb{E}_{it} \left[1\{x_{it+1}(1) = -1\} - 1\{x_{it+1}(0) = -1\} \right] - \widetilde{\varepsilon}_{it} \quad (7)$$

with $\widetilde{\alpha} \equiv \alpha(1) - \alpha(0)$, and $-\widetilde{\varepsilon}_{it} \equiv \varepsilon_{it}(1) - \varepsilon_{it}(0)$, and with $\mathbb{E}_{it}(.)$ representing the expectation operator conditional on the individual's information at t. More specifically, we have that:

$$\mathbb{E}_{it} \left[n_{it}^{(-1)} + n_{it}^{(0)} \right] = \widetilde{n}_{it}^{(-1,0)} \; [1 - Q_t] \quad (8)$$

with $\widetilde{n}_{it}^{(-1,0)} \equiv \sum_{j \in \mathcal{C}_i, j \neq i} 1\{\widetilde{x}_{jt} = [-1 \; or \; 0]\}$.

The term $\mathbb{E}_{it} \left[1\{x_{it+1}(1) = -1\} - 1\{x_{it+1}(0) = -1\} \right]$ deserves more detailed explanation. A necessary condition to reach state $x_{it+1} = -1$ is that $x_{it} = 0$ and $a_{it} = 0$. Therefore, the indicator $1\{x_{it+1}(1) = -1\}$ is equal to zero. Furthermore, the expectation $\mathbb{E}_{it}[1\{x_{it+1}(0) = -1\}]$ is equal to the probability that $x_{it} = 0$ times the probability of infection conditional on ($x_{it} = 0$ and $a_{it} = 0$):

$$\mathbb{E}_{it}[1\{x_{it+1}(0) = -1\}] \quad = \quad \frac{S_t(0)}{S_t(0) + S_t(-1)} \; \Lambda\left(\eta_0 + \eta_1 \; \mathbb{E}_{it}\left[n_{it}^{(-1)}\right]\right) \tag{9}$$

and

$$\mathbb{E}_{it}\left[n_{it}^{(-1)}\right] = \widetilde{n}_{it}^{(-1,0)} \frac{S_t(-1)}{S_t(0) + S_t(-1)}[1 - Q_t] \tag{10}$$

We can use the following more compact notation to represent the difference between the expected utilities:

$$u(\{\widetilde{x}_{jt}\}_{j \in \mathcal{C}_i, j \neq i}, Q_t, \mathbf{S}_t)'\theta - \widetilde{\varepsilon}_{it} \tag{11}$$

where θ is the vector of structural parameters in preferences, $(\widetilde{\alpha}, \beta, \delta)'$, and $u(\{\widetilde{x}_{jt}\}_{j \in \mathcal{C}_i, j \neq i}, Q_t, \mathbf{S}_t)$ is the vector $(1, -\mathbb{E}_{it}[n_{it}^{(-1)} + n_{it}^{(0)}] - n_{it}^{(1)}, \mathbb{E}_{it}[1\{x_{it+1}(0) = -1\}])$.

The best response probability function of individual i at period t is:

$$P_{it} = \Lambda\left(\; u(\{\widetilde{x}_{jt}\}_{j \in \mathcal{C}_i, j \neq i}, Q_t, \mathbf{S}_t)'\theta \; \right) \tag{12}$$

1.2.2 Equilibrium conditions

Given an initial condition $\mathbf{x}_0$, an equilibrium is a sequence of probabilities $\{Q_t, \mathbf{S}_t : t \geq 0\}$ that satisfy the following conditions at every period t:

$$\begin{array}{llcl}
(i) & P_{it} & = & \Lambda\left(\; u(\{\widetilde{x}_{jt}\}_{j \in \mathcal{C}, j \neq i}, Q_t, \mathbf{S}_t)'\theta \; \right) \quad \text{for any } i \in \mathcal{I} \\
(ii) & Q_t & = & \dfrac{\sum_{i \in \mathcal{I}} 1\{\widetilde{x}_{it} = [-1 \; or \; 0]\} \; P_{it}}{\sum_{i \in \mathcal{I}} 1\{\widetilde{x}_{it} = [-1 \; or \; 0]\}} \\
(iii) & \Pi_t^v & = & \dfrac{\sum_{i \in \mathcal{I}} 1\{\widetilde{x}_{it} = [-1 \; or \; 0]\} \; \Lambda\left(\eta_0 + \eta_1 \; \mathbb{E}_{it}\left[n_{it}^{(-1)}\right]\right)}{\sum_{i \in \mathcal{I}} 1\{\widetilde{x}_{it} = [-1 \; or \; 0]\}} \\
(iv) & S_t(-3) & = & S_{t-1}(-3) + \pi^m \; S_{t-1}(-2) \\
(v) & S_t(1) & = & S_{t-1}(1) + (1 - \pi^m) \; S_{t-1}(-2) \\
(vi) & S_t(-2) & = & S_{t-1}(-1) \\
(vii) & S_t(-1) & = & S_{t-1}(0) \; [1 - Q_{t-1}] \; \Pi_{t-1}^v \\
(viii) & S_t(0) & = & 1 - S_t(-3) - S_t(-2) - S_t(-1) - S_t(1)
\end{array} \tag{13}$$

1.3 Estimation

We consider the structural estimation of this model under two different types of datasets.

Micro-level data. It is a random sample of N individuals over T periods of time. The number of individuals N is large (asymptotics in N). The researcher observes a_{it}, x_{it}, and $\{a_{jt}, x_{jt} : j \in C_i\}$ for every individual i and period t in the sample.

Aggregate-level data. It is a random sample of M societies over T periods of time. The number of societies M is large (asymptotics in M). The researcher observes $\{Q_{mt}, S_{mt}(x) : x \in \mathcal{X}\}$ for every society m and period t in the sample.

Note that the researcher observes the state x_{it} and not only $\widetilde{x}_{it}$. This is because, under the conditions of the model (i.e., in particular, the assumption that any infected individual always develops symptoms), we can infer whether $x_{it} = 0$ or $x_{it} = -1$ by observing the value of $\widetilde{x}_{i,t+1}$. That is, given that we know $\widetilde{x}_{it} = [-1 \text{ or } 0]$ and $\widetilde{x}_{i,t+1}$, we have that: $x_{it} = 0$ if and only if $\widetilde{x}_{i,t+1} = [-1 \text{ or } 0]$; and $x_{it} = -1$ if and only if $\widetilde{x}_{i,t+1} = -2$.

We are interested in using these data to estimate the vector of structural parameters of the model: $\widetilde{\alpha}$, β, δ, π^m, η_0, and η_1.

---

2 Questions

QUESTION 1 [30 POINTS]. Properties of the model

1.1. [10 points] Consider the best response probability function in equation (12).

(a) Obtain the expression for the derivative $\dfrac{\partial P_{it}}{\partial Q_t}$.

(b) Given that $\beta > 0$, $\delta > 0$, and $\eta_1 > 0$, discuss the sign of this derivative.

(c) Are individuals confinement decisions strategic complements or substitutes? Under which conditions?

1.2. [10 points] Consider the equilibrium condition (ii) in equation (13).

(a) Show that this equation defines a fixed point in Q_t.

(b) Prove that there is at least one value of Q_t that solves this fixed point problem.

(c) Based on your answer to Questions 1.1(b) and 1.1(c), impose sufficient conditions on the primitives such that $\dfrac{\partial P_{it}}{\partial Q_t} < 0$. Under these conditions, prove that fixed point Q_t is unique.

(d) Based on your answer to Questions 1.1(b) and 1.1(c), impose sufficient conditions on the primitives such that $\dfrac{\partial P_{it}}{\partial Q_t} > 0$. Under these conditions, show that this model is a coordination games with multiple equilibria for Q_t: there is a "good" equilibria with high Q, and a "bad" equilibria with low Q.

(d) Propose an algorithm to compute an equilibrium Q_t (given that the researcher knows all the primitives an structural parameters).

1.2. [**10 points**] Consider all the equilibrium conditions in equation (13).

(a) Describe in detail an algorithm to compute an equilibrium $\{Q_t, \mathbf{S}_t : t \geq 0\}$.

QUESTION 2 [30 POINTS]. Estimation of the model

2.1. [**15 points**] Consider the micro-level dataset. Propose an estimation method to estimate all the structural parameters. Describe this method in detail.

2.2. [**15 points**] Consider the aggregate-level dataset. Propose an estimation method to estimate all the structural parameters. Describe this method in detail.

QUESTION 3 [20 POINTS]. Counterfactual policy experiments

Suppose that the government is interested in implementing mandatory confinement of all the individuals. More specifically, it imposes an amount of penalty of W "utils" that should be paid by those individuals in state $\widetilde{x} \in \{-1, 0\}$ who choose not to follow confinement.

3.1. [**10 points**] Explain how to obtain the counterfactual equilibrium and how to measure the effects of this equilibrium.

3.2. [**5 points**] Perform welfare analysis. Measure the welfare effects of this counterfactual policy if implemented at some period $t^* \geq 1$. To implement this welfare analysis, assume that the utility function for individuals at states $x_{it} = -2$ and $x_{it} = 1$ is the same as the utility that we have specified for individuals at states $x_{it} = -1$ and 0 [though individuals at $x_{it} = -2$ are forced to choose $a_{it} = 1$ and individuals at $x_{it} = 1$ always choose $a_{it} = 0$]. For individuals at state $x_{it} - 3$ (death), consider that the utility is equal to a negative value $-D$.

3.3. [**5 points**] Explain the tradeoff in the choice of the value of the implementation period t^*, and how to choose t^* optimally.

QUESTION 4 [20 POINTS]. Coding an equilibrium of the model.

Consider a society that consists of $I = 100$ indexed by $i = 1, 2, ..., 100$. For any individual $2 \leq i \leq 99$, the set of social contacts is $\mathcal{C}_i = \{i - 1, i + 1\}$. For individuals 1 and 100, we have that $\mathcal{C}_1 = \{100, 2\}$ and $\mathcal{C}_{100} = \{99, 1\}$.

4.1. [**10 points**] Write code that computes and equilibrium of the model. Choose values of structural parameters $\widetilde{\alpha}$, β, δ, π^m, η_0, and η_1 and implement this code.

4.2. [**10 points**] Choose an initial condition an simulate the equilibrium. Present figures with the time series for Q_t and $\mathbf{S}_t$.

END OF THE PROBLEM SET