Research Proposal:
Robustness of Distributional Assumption on Structural Error in Moment Inequality Estimation

Shuang Wang

September 19, 2018

1   Introduction

In various empirical problems, such as entry games, discrete choice with social interactions and network formation, economic theory are stranded in selection among multiple equilibria. Take a 2×2 entry game as an example, when one firm's entry has negative effects on the other's, the equilibrium could be either firm enters as a monopolist. As a result, without further assumptions on equilibrium selection mechanism or restriction on heterogeneity across agents, it is infeasible to back out the parameters in the profit function with the method of maximum likelihood estimation (MLE).

In the recent decade, moment inequality emerges as a method to estimate such *incomplete* models. As summarized in **?**, there are mainly two different ways of constructing inequalities that are based on different assumptions on *measurement errors* and *structural errors*.

The first approach is labeled as the *generalized discrete choice* approach (GDC henceforth), as it generalizes familiar discrete choice model to allow for multiple-agent interaction. The ideas behind this approach date to **?**, and are developed in more details in papers by **?** and **?**. The basic idea of GDC is constructing upper and lower bounds for the probability of each observed outcome, then all parameters that makes the bounds cover the observed probabilities are included in the parameter sets. However, GDC sometimes are undesirable due to computation burden and the risk of misspecification. Specifically, since there are usually no analytical forms for the probability bounds, researchers have to use simulated ones instead. The fact that the number of the possible outcomes increases exponentially with the number of players may induce a computational issue. In addition, the inevitable distributional assumption on the unobservables is likely to expose the model to misspecification. **?** has shown that the normal distribution assumption on the *structural error* leads to biased estimation.

The second approach, *profit inequality* (PI henceforth), is based on the idea of *revealed preference* in the sense that inequalities are generated by the assumption that expected profits from the observed choice must be at least as great as those from alternative feasible choices. Since it avoids the computation of probabilities, those two issues of GDC becomes trivial. However, unless the structural errors are arbitrarily assumed away or in a restricted form, PI requires a weight function to circumvent the them without violating the inequalities, or there would be *selection bias* problem. And such a weight function may be infeasible in a specific cases, for example, the entry game with additive structural errors as in in **?**.

Therefore, the objective of this paper is using the method of PI to reestimate the model in **?** where GDC is used originally, checking whether the estimates are robust to changes in the distributional assumptions. Furthermore, since the weight function required by PI is infeasible in this specific case, I need to find a new approach to address the selection bias problem.

2   Related Literature

This paper aims at contributing to the following two strands of literature.

First, I am trying to find a new approach to resolve the selection bias problem in the applications of PI so that it allows agent-specific structural errors and interaction among a relative large number of agents with a reasonable computational burden. In particular, although PI does not require distributional assumption on the unobservable part of the profit function, it does require a nonnegative weight function that makes the weighted sum of the unobservables negligible in the inequalities. i.e. if the unobservables have a expectation with a sign opposite to the inequality, ignoring them will not change the sign of the inequality.

Such a requirement would be trivial if the unobservables are *measurement errors* or *structural errors independent* to agents' decisions with a zero expectation. For example, in **?**, the marginal ATM installation cost of each bank has an idiosyncratic deviation from the average level, and it is unobservable to researchers. However, it is assumed to be invariant with the number of ATM installed by the bank, so its sample analog across all banks in the market goes to zero in probability. Similar approach is adopted in **?**. On the other hand, when the *structural errors* are decision-specific, such a weight function may be infeasible. For instance, in the entry game among airlines, there is an unobserved component linearly added to the profit function, so it is likely to be relatively higher for the entrants who yield nonnegative profits in the market. Such a positive correlation is the source of *selection bias* in moment inequality estimation. And ignoring them may turn the inequality over. More details can be found in Section 3.

In the existing literature, the most commonly-used approach to address this *selection bias* problem is to assume that the unobservables is some market-level fixed effects, so the unobservables can be canceled out by differencing the inequalities of two agents in the same market.(**?**, **?**). As far as I know, the only work that allows for agent-level, decision-dependent unobservables is **?**. The author estimates the fixed cost for each PC makers to offer a new configuration, and there is a mean-zero configuration-specific cost shock that are unobservable to econometricians. To construct... the unlisted products is assumed to be bounded above. It is plausible in that specific context because potential products considered there are CPU selected from a finite set of options installed in existing notebook PC product lines. However, it is inapplicable for the entry in airline markets, since the decision of not entering might have been characterized by an infinitely-large negative demand shock. That is why we need a new approach to solve the selection bias problem.

Second, to my knowledge, there is no existing works that compare the results from GDC and PI with real data from entry games. The most closely-related work is **?**. The author lists different assumptions of GDC and PI and checks their robustness with simulated data and the real data from a network formation problem in **?**. The result indicates that the estimates from both models are surprisingly robust to all likely sources of problems but one: the need to assume a distribution for the GDC model. This on the other hand indicates reestimating **?** with PI might be a challenge to the results from GDC model.

3 Model and Assumptions

In order to make comparison to the results in **?**, I will follow their basic settings and discuss the implications of GDC and PI assumptions in this specific case.

The main idea of **?** is estimating the heterogeneous competitive effects among airlines, i.e., how would one firm's entry affects the profit of other firms. Specifically, suppose there are K potential airlines indexed by $i = 1, ..., K$ in market m, and a market is defined as a trip between two airports. Let $\mathcal{J}_{im}$ denote the information set available to airline i when decision is made for market m. Its profit (π_{im}) is determined by its own entry decision($y_{im} \in \{0, 1\}$), other airlines' decisions ($y_{-im} = \{y_{jm}\}_{j \neq i}$) and an additional set of airline and market characteristics (X_{im}), i.e.

$$
\begin{aligned}
\pi(y_{im}, y_{-im}, X_{im}) &= y_{im}(S'_m \alpha_i + Z'_{im} \beta_i + W'_{im} \gamma_i + \sum_{j \neq i} \delta^i_j y_{jm} + \sum_{j \neq i} Z'_{jm} \phi^i_j y_{jm} + \epsilon_{im}) \\
&= y_{im} S'_m \alpha_i + Z'_{im} \beta_i + W'_{im} \gamma_i + \sum_{j \neq i} \delta^i_j y_{jm} + \sum_{j \neq i} Z'_{jm} \phi^i_j y_{jm} + y_{im} \epsilon_{im} \quad (1)
\end{aligned}
$$

where S_m are market-specific explanatory variables, $Z_m = (Z_{1m}, ..., Z_{Km})$ are firm-specific explanatory variables that enter the profit functions of all firms and W_{im} are firm-specific variables that enter the profit function of firm i only. They are all components of the set of observable airline and market characteristics, i.e. $X^o_{im} = [S'_m, Z'_{im}, W'_{im}]$ and $X^o_{im} \in X_{im}$. Respectively, ϵ_{im} indicates the airline and market characteristics that are unobservable to researchers but observable to agents when the entry decisions are made.

3.1 Common Assumptions

For both GDC and PI, we need to assume that airlines make decisions according to their expected profit conditional on their available information, i.e.,

Assumption 1: We assume

$$E[\pi(\sigma(\mathcal{J}_{im}), \mathbf{y}_{-im}, \mathbf{X}_{im})|\mathcal{J}_{im}, \mathbf{y}_{im} = \sigma(\mathcal{J}_{im})] \geq \sup_{y' \in \{0,1\}} E[\pi(y', \mathbf{y}_{-im}, \mathbf{X}_{im})|\mathcal{J}_{im}, \mathbf{y}_{im} = y']^1,$$

where function $\sigma(\cdot)$ gives airline i's strategy given its information set $\mathcal{J}_{im}$, $\mathbf{y}_{-im}$ is the vector of other competitors' strategies which is a random variable to airline i and $\mathbf{X}_{im}$ includes other random characteristics(I use boldface to distinguish random variables from their realization). The expectation is over both $\mathbf{y}_{-im}$ and $\mathbf{X}_{im}$. Assumption 1 is a necessary condition for any Bayes-Nash equilibrium. It implies that given agent i's belief about its competitors' strategies and market conditions, the realized action must be at least as profitable as any other available counterfactual choices.

In reality, the conditional distributions of both $\mathbf{y}_{-im}$ and $\mathbf{X}_{im}$ could change with y_{im}. For instance, say market price is a component in $\mathbf{X}_{im}$, if airline i decides to enter market m, the increased competition may lower the market price, which may further drive those less profitable airlines out of market thus change $\mathbf{y}_{-im}$. In this respect, we need a "structural" model which characterizes such interaction for future counterfactual analysis. However, to avoid such complexity, we assume that agent's perception about other agents' strategies and market characteristics does not change with its own actions given that the information set is kept unchanged, i.e.,

Assumption 2: The distribution of $(\mathbf{X}_{im}, \mathbf{y}_{-im})$ conditional on $\mathcal{J}_{im}$ and $\mathbf{y}_{im} = y$ does not depend on y.

With *Assumption 2*, *Assumption 1* can be rewritten without conditional on $\mathbf{y}_{im}$,

$$E[\pi(\sigma(\mathcal{J}_{im}), \mathbf{y}_{-im}, \mathbf{X}_{im})|\mathcal{J}_{im}] \geq \sup_{y' \in \{0,1\}} E[\pi(y', \mathbf{y}_{-im}, \mathbf{X}_{im})|\mathcal{J}_{im}] \tag{2}$$

Next, in accordance with **?**, we further assume agents have complete information,

Assumption 3: $\forall y_{im} \in \{0, 1\}$

$$E[\pi(y_{im}, \mathbf{y}_{-im}, \mathbf{X}_{im})|\mathcal{J}_{im}] = \pi(y_{im}, y_{-im}, X_{im}),$$

To rationalize the observed data, it is necessary to have errors in profit or utility function, either expectational error or structural error. In theory, structural error is always there as long as it is agreed that econometricians can not observe everything that agents observe. Whether to have expectational error depends on information assumption, complete or incomplete.

so there is no uncertainty in either firm and market characteristics ($\mathbf{X}_{im}$) or in the strategies of the firm's competitors ($\mathbf{y}_{-im}$).

Generally, PI does not require complete information, it allows agents to have expectational errors and denote it as $v^1_{i,m,y}$, where $v^1_{i,m,y} = \pi(y_{im}, y_{-im}, X_{im}) - E[\pi(y_{im}, \mathbf{y}_{-im}, \mathbf{X}_{im})|\mathcal{J}_{im}]$. However, since its unconditional expectation is 0 by definition, such an assumption is not too restrictive in the estimation.

As a result, Eq.(**??**) can be further simplified by *Assumption 3* as

$$\pi(\sigma(\mathcal{J}_{im}), y_{-im}, X_{im}) \geq \sup_{y' \in \{0,1\}} \pi(y', y_{-im}, X_{im}) \tag{3}$$

3.2 Distinguishing Assumptions

The assumptions of GDC and PI diverges regarding the *structural error* generated by the difference between the real profit function $\pi(y_{im}, y_{-im}, X_{im})$ and researcher's approximation $r(y_{im}, y_{-im}, X^o_{im}; \theta)$. Specifically, since researchers usually do not have much knowledge about the real functional form of $\pi(\cdot)$ and cannot observe the whole vector of X_{im}, they instead assume an approximate profit function $r(\cdot)$ which is a function of the observed subset of characteristics (X^o_{im}) and is identified up to a finite set of parameters θ. Following **?**, we define such structural error as

$$v^2_{i,m,y} = \pi(y_{im}, y_{-im}, X_{im}) - r(y_{im}, y_{-im}, X^o_{im}, \theta). \tag{4}$$

---

[1] Throughout this proposal, we use $E[\cdot]$ to denote expectation with respect to the true data generating process. In general, agents' expectations are allowed to deviated from the "true expectation". The "subjective expectation" is usually denoted as $\mathcal{E}[\cdot]$ in literature(**?**) and additional assumptions are needed for estimation. Here, for simplicity, we assume "correct expectation", following **?**, i.e., $E[\cdot] = \mathcal{E}[\cdot]$

Therefore, in Eq.(**??**),

$$r(y_{im}, y_{-im}, X^o_{im}, \theta) = y_{im}(S'_m \alpha_i + Z'_{im}\beta_i + W'_{im}\gamma_i + \sum_{j \neq i} \delta^i_j y_{jm} + \sum_{j \neq i} Z'_{jm}\phi^i_j y_{jm}) \tag{5}$$

$$v^2_{i,m,y} = y_{im}\epsilon_{im} \tag{6}$$

In **?**, in order to solve the equilibria in each market, they made distributional assumptions on ϵ_{im} so that they can back out the "real" profit $\pi(\cdot)$ by simulation. Specifically,

Assumption GDC: We assume

$$\epsilon_{im} = u_m + u^o_m + u^d_m + u_{im},$$

where u_m is market-specific unobserved heterogeneity, u^o_m is an error that is common across all markets whose origin is o, u^d_m is an error that is common across all markets whose origin is d and u_{im} is firm-specific unobserved heterogeneity. u_m, u^o_m, u^d_m and u_{im} are *independent and normally* distributed.

With such an assumption, **?** use simulation to calculate the equilibira where the condition in Eq. (**??**) is satisfied for all firms in the market. Explicitly, for each draw of $\{\epsilon_{im}\}_{i,m}$, if $\mathbf{y}_m = \{y_{im}, y_{-im}\}$ is the unique equilibrium, they add one to both the upper and lower bounds for the probability of $\mathbf{y}_m$; if it is one of the multiple equilibria, they add one to the upper bound only, then they take the average over all random draws. Then those parameter values that make the inequalities hold are included in the identified set.

Rather than constructing bounds for outcome probabilities, **?** directly uses Eq.(**??**) for estimation. Since for any airline i, the profit is zero if it chooses not to enter the market, i.e., $\pi(0, y_{-im}, X_{im}) = 0$, for any observed $(\{y_{im}\}_i, \{X^o_{im}\}_i)$[2], we have

$$y_{im} = 1 \quad \text{iff} \quad \pi(1, y_{-im}, X_{im}) \geq 0$$

which is equivalent to

$$(-1)^{y_{im}+1}[r(1, y_{-im}, X^o_{im}, \theta) + \epsilon_{im}] \geq 0 \tag{7}$$

However, ϵ_{im} is unobservable to researchers. To circumvent it in estimation, **?** makes the following assumption.

Assumption PI: Let $h(y_{im}, y_{-im}, X^o_{im})$ be a nonnegative function whose value can depend on market structure $\{y_{im}, y_{-im}\}$ and the observed firm and market characteristics X^o_{im}. Assume that

$$E[\sum_i h(y_{im}, y_{-im}, X^o_{im})(-1)^{y_{im}+1}(\epsilon_{im})] \tag{8}$$

Assumption PI implies that although ϵ_{im} is not observable to econometricians, it is ignorable in limit when the number of markets goes to infinity so that we could focus on $r(\cdot)$ in estimation.

4 Issues in applying PI

4.1 Selection Bias

As Eq.(**??**) indicates, the difference in structural error $\Delta v^2_{i,m,y} = (-1)^{y_{im}+1}\epsilon_{im}$ depends on firm's decision y_{im}, which makes it a more complicated case than **?**. In this section, we will show that due to the *sample selection* problem, *Assumption PI* is infeasible in this specific context of **?**, i.e. the required weight function $h(\cdot)$ does not exist. Specifically, define I_m as the set of incumbents in market m, i.e.,$I_m = \{k : y_{km} = 1\}$. Then Eq.(**??**) in *Assumption PI* can be rewritten as

---

[2]Note that we can not say anything about unobserved outcomes. For example, in a market with 2 potential entrants, if $(1, 0)$ is observed, it does not necessary indicate that $\pi_i(0, 1)$ is less than 0

$$E[\sum_i h(y_{im}, y_{-im}, X_{im}^o)(-1)^{y_{im}+1}\epsilon_{im}] = E[\sum_{j\in I_m} h(1, y_{-jm}, X_{jm}^o)\epsilon_{jm}|y_{jm}=1] - E[\sum_{k\notin I_m} h(0, y_{-km}, X_{km}^o)\epsilon_{km}|y_{km}=0]$$

$$= \sum_{j\in I_m} h(1, y_{-jm}, X_{jm}^o)E[\epsilon_{jm}|y_{jm}=1] - \sum_{k\notin I_m} h(0, y_{-km}, X_{km}^o)E[\epsilon_{km}|y_{km}=0]$$

Consider

$$E[\epsilon_{jm}|y_{jm}=1] = E[\epsilon_{jm}|r(1, y_{-jm}, X_{jm}^o, \theta) + \epsilon_{jm} > 0]$$
$$= E[\epsilon_{jm}|\epsilon_{jm} > -r(1, y_{-jm}, X_{jm}^o, \theta)]$$
$$E[\epsilon_{km}|y_{km}=0] = E[\epsilon_{km}|\epsilon_{km} < -r(0, y_{-km}, X_{km}^o, \theta)]$$

Then, given $E[\epsilon_{im}] = 0$, $E[\epsilon_{jm}|y_{jm}=1] > 0$ and $E[\epsilon_{km}|y_{km}=0] < 0$. Therefore, there is no nonnegative $h(y_{im}, y_{-im}, X_{im}^o)$ that makes their weighted sum to be less than 0.

Furthermore, since ϵ_{im} is firm-and-market-specific, it is impossible to eliminate it by differencing two inequalities with the same ϵ as is some preceding studies. Therefore, I need a new approach to solve this selection bias problem.

5 Preliminary Empirical Strategy

5.1 A starting point - PI estimation w/o structural errors:

To start the analysis, we first circumvent the complexity of selection bias by assuming away the structural error $v_{i,m,y}^2$. This exercise would be useful in checking how identified set changes with assumptions. Instead, we allow for a measurement error in *cost*, which is an element of W_{im}' in **?**. In general, we would expect the identified set from PI without distributional assumptions to be consistent, and cover that from GDC. Therefore, if this not the case, we should be worried about the results from GDC to be inconsistent. Furthermore, we could do the comparison with simulated data where the true identified sets are know, as in **?**.

The profit function in this scenario is

$$\pi(y_{im}, y_{-im}, X_{im}) = y_{im}(S_m'\alpha_i + Z_{im}'\beta_i + W_{im}'\gamma_i + \sum_{j\neq i}\delta_j^i y_{jm} + \sum_{j\neq i}Z_{jm}'\phi_j^i y_{jm})$$

$$= y_{im}(S_m'\alpha_i + Z_{im}'\beta_i + \sum_{j\neq i}\delta_j^i y_{jm} + \sum_{j\neq i}Z_{jm}'\phi_j^i y_{jm}) + y_{im}W_{im}'\gamma_i$$

$$= y_{im}(S_m'\alpha_i + Z_{im}'\beta_i + \sum_{j\neq i}\delta_j^i y_{jm} + \sum_{j\neq i}Z_{jm}'\phi_j^i y_{jm}) + (y_{im}W_{im}'^o + \varphi_{im})\gamma_i$$

$$= y_{im}(S_m'\alpha_i + Z_{im}'\beta_i + W_{im}'^o\gamma_i + \sum_{j\neq i}\delta_j^i y_{jm} + \sum_{j\neq i}Z_{jm}'\phi_j^i y_{jm}) + \varphi_{im}\gamma_i$$

$$= r(y_{im}, y_{-im}, X_{im}^o, \theta) + v_{i,m}^1 \tag{9}$$

Assume that $E[\varphi_{im}|\mathcal{J}_{im}] = 0$, i.e., $E[v_{i,m}^1|\mathcal{J}_{im}] = 0$.

Furthermore, we begin with model (1) with the restrictions $\phi_j^i = \phi_j = 0$, $\beta_i = \beta$, $\alpha_i = \alpha$ and $\delta_j^i = \delta_j$, $\forall i, j$, then

$$r(y_{im}, y_{-im}, X_{im}^o, \theta) = y_{im}(S_m'\alpha + Z_{im}'\beta + W_{im}'^o\gamma + \sum_{j\neq i}\delta_j y_{jm})$$

Then, Eq.(**??**) can be rewritten as

$$(-1)^{y_{im}+1}(r(1, y_{-im}, X_{im}^o, \theta) + v_{i,m}^1) \geq 0 \tag{10}$$

Taking expectations and using the fact that $E[v_{i,m}^1|\mathcal{J}_{im}] = 0$, we find that

$$E[(-1)^{y_{im}+1}r(1, y_{-im}, X_{im}^o, \theta)|X_{im}^o \in \mathcal{J}_i m] \geq 0 \tag{11}$$

Translating expectations into sample means, the equations for estimation is therefore

$$\frac{1}{MI} \sum_m \sum_i (-1)^{y_{im}+1} r(1, y_{-im}, X_{im}^o, \theta) \geq 0 \tag{12}$$

And the objective function is

$$Q_n(\theta) = \sum_m \sum_i \|(-1)^{y_{im}+1} r(1, y_{-im}, X_{im}^o, \theta)\|_+$$

where $\|a\|_+ = a1[a \leq 0]$.

Then we can construct the confidence sets for θ following **?** as **?** does.

5.2 A semiparametric first-stage - PI estimation w/ structural errors

Similarly, we begin with the simplest specification of the profit function where $\alpha_i = \alpha$, $\beta_i = \beta$, $\gamma_i = \gamma$, $\delta_j^i = \delta_j$ and $\phi_j^i = 0$ for $\forall i, j$, i.e.

$$\pi(y_{im}, y_{-im}, X_{im}) = y_{im}(S_m'\alpha + Z_{im}'\beta + W_{im}'\gamma + \sum_{j \neq i} \delta_j y_{jm}) + y_{im}\epsilon_{im} \tag{13}$$

Assumptions: Nash Equilibrium + Complete information + Negative competition effect + ***iid* shock**

The main impediment to MLE is multiple equilibrium, however, those markets where all firms enter or no firms enters are immune to such a complexion. Taking those markets having all firms entered as an example, given the Nash equilibrium definition and negative competition effects assumption, if the firm can benefit from entry even though all other firms enter, it will still enter if any other firm exit. Therefore, any market structure with one or more firms outside the market can not be an equilibrium. In other words, $\pi(y_{im} = 1, y_{-im} = 1, X_{im}) > 0$, $\forall i$ and $\pi(y_{im} = 0, y_{-im} = 0, X_{im}) > 0$, $\forall i$ are both necessary and sufficient conditions, which enables us to write down the likelihood.

$$
\begin{aligned}
Pr(0, 0, ..., 0|\mathbf{X}_m^o) &= Pr(\pi(y_{im} = 0, y_{-im} = 0, X_{im}) > 0, \ \forall i) \\
&= Pr(\epsilon_{im} < -(S_m'\alpha + Z_{im}'\beta + W_{im}'\gamma), \ \forall i) \\
&= \prod_i \{1 - F(S_m'\alpha + Z_{im}'\beta + W_{im}'\gamma)\}
\end{aligned}
$$

$$
\begin{aligned}
Pr(1, 1, ..., 1|\mathbf{X}_m^o) &= Pr(\pi(y_{im} = 1, y_{-im} = 1, X_{im}) > 0, \ \forall i) \\
&= Pr(\epsilon_{im} > -(S_m'\alpha + Z_{im}'\beta + W_{im}'\gamma + \sum_{j \neq i} \delta_j), \ \forall i) \\
&= \prod_i F(S_m'\alpha + Z_{im}'\beta + W_{im}'\gamma + \sum_{j \neq i} \delta_j)
\end{aligned}
$$

where $F(\cdot)$ is the distribution of ϵ_{im}, $\forall i, m$.

Next, we will focus two types of markets, markets with all firm entered and those with no firms entered, construct a nonparametric estimate of $F(\cdot)$ for a given value of θ, plug the estimated cdf $\hat{F}(\cdot)$ into the moment inequalities, then any values of θ yielding the moment inequalities are in the identified set.

Nonparametric $\hat{F}(\cdot)$

The comparative probability ratio between $(0, 0, ..., 0)$ and $(1, 1, ..., 1)$ is determined by the index in $F(\cdot)$ and $F(\cdot)$ itself. Therefore, intuitively, for any given set of θ, there is some optimal $\hat{F}(\cdot)$ makes the predicted ratio of conditional log likelihood

$$\frac{\sum_i log\{1 - \hat{F}(S_m'\alpha + Z_{im}'\beta + W_{im}'\gamma)\}}{\sum_i log\hat{F}(S_m'\alpha + Z_{im}'\beta + W_{im}'\gamma + \sum_{j \neq i} \delta_j)}$$

"match" the observed one [3] [4]

$$\frac{logPr(0, 0, ..., 0|\mathbf{X}_m^o)}{logPr(1, 1, ..., 1|\mathbf{X}_m^o)}$$

---

[3]In **?**, the use the nonparametric cdf to find the optimal θ by MLE; however, we can only construct the likelihood of 2 possible outcomes using $\hat{F}$ and the θ that maximizes such a conditional likelihood is not the interest of this paper.

[4]Such a method might be rationalized by "IIA"

Moment Inequality

Then, we will construct moment inequalities and use the empirical $\hat{F}(\cdot)$ to control selection bias.

Taking each market as an observation, we have

$$
\begin{aligned}
\sum_{i \in I_m} r(1, y_{-im}, X_{im}^o, \theta) + \epsilon_{im} &\geq 0 \\
\sum_{i \notin I_m} r(1, y_{-im}, X_{im}^o, \theta) + \epsilon_{im} &\leq 0
\end{aligned}
\tag{14}
$$

which necessarily implies the following moment inequality

$$
\begin{aligned}
E[\sum_{i \in I_m} r(1, y_{-im}, X_{im}^o, \theta) + \epsilon_{im} | y_{im}, y_{-im}, X_m^o, \mathcal{S}] &\geq 0 \\
E[\sum_{i \notin I_m} r(1, y_{-im}, X_{im}^o, \theta) + \epsilon_{im} | y_{im}, y_{-im}, X_m^o, \mathcal{S}] &\leq 0
\end{aligned}
$$

where (y_{im}, y_{-im}, X_m^o) represents an observed market, and $\mathcal{S}$ is some unobservable equilibrium selection mechanism that is uniform across markets.[5] Since the deterministic part of the profit $r(\cdot)$ does not depend on the selection mechanism, the conditional expectations above can be further rewritten as

$$
\begin{aligned}
E[\sum_{i \in I_m} r(1, y_{-im}, X_{im}^o, \theta) | y_{im}, y_{-im}, X_m^o] + E[\sum_{i \in I_m} \epsilon_{im} | y_{im}, y_{-im}, X_m^o, \mathcal{S}] &\geq 0 \\
E[\sum_{i \notin I_m} r(1, y_{-im}, X_{im}^o, \theta) | y_{im}, y_{-im}, X_m^o] + E[\sum_{i \notin I_m} \epsilon_{im} | y_{im}, y_{-im}, X_m^o, \mathcal{S}] &\geq 0
\end{aligned}
\tag{15}
$$

In particular, $E[\epsilon_{im} | y_{im}, y_{-im}, X_m^o, \mathcal{S}]$ is infeasible to calculate without assumption on equilibrium selection mechanism. To circumvent those arbitrary assumptions, we instead use an upper bound for that conditional expectation when firm i is an incumbent, and a lower bound for when the firm is a potential entrant i.e.,

$$
\begin{aligned}
E[\epsilon_{im} | \epsilon_{im} \geq - \min_{y_{-im}} r(1, y_{-im}, X_{jm}^o, \theta), X_{im}^o] &\geq E[\epsilon_{im} | y_{im} = 1, y_{-im}, X_m^o, \mathcal{S}], \ \forall i \in I_m \\
E[\epsilon_{im} | \epsilon_{im} \leq - \max_{y_{-im}} r(1, y_{-im}, X_{jm}^o, \theta), X_{im}^o] &\leq E[\epsilon_{im} | y_{im} = 0, y_{-im}, X_m^o, \mathcal{S}], \ \forall i \notin I_m
\end{aligned}
\tag{16}
$$

Therefore, replacing the conditional expectation of ϵ_{im} in Eq.(**??**) with their bounds would give us

$$
\begin{aligned}
E[\sum_{i \in I_m} r(1, y_{-im}, X_{im}^o, \theta) | y_{im}, y_{-im}, X_m^o] + E[\sum_{i \in I_m} \epsilon_{im} | \epsilon_{im} \geq - \min_{y_{-im}} r(1, y_{-im}, X_{jm}^o, \theta), X_m^o] &\geq 0 \\
E[\sum_{i \notin I_m} r(1, y_{-im}, X_{im}^o, \theta) | y_{im}, y_{-im}, X_m^o] + E[\sum_{i \notin I_m} \epsilon_{im} | \epsilon_{im} \leq - \max_{y_{-im}} r(1, y_{-im}, X_{jm}^o, \theta), X_m^o] &\geq 0
\end{aligned}
\tag{17}
$$

Finally, we will construct sample analog for the following conditional moment,

$$
\begin{aligned}
E[r(X_m^o) \times (\sum_{i \in I_m} r(1, y_{-im}, X_{im}^o, \theta) + E[\epsilon_{im} | \epsilon_{im} \geq - \min_{y_{-im}} r(1, y_{-im}, X_{jm}^o, \theta), X_m^o]) | y_{im}, y_{-im}, \mathcal{S}] &\geq 0 \\
E[r(X_m^o) \times (\sum_{i \in I_m} r(1, y_{-im}, X_{im}^o, \theta) + E[\epsilon_{im} | \epsilon_{im} \leq - \max_{y_{-im}} r(1, y_{-im}, X_{jm}^o, \theta), X_m^o]) | y_{im}, y_{-im}, \mathcal{S}] &\leq 0
\end{aligned}
$$

6 Data

In **?**, the main data comes from the second quarter of the 2001 Airline Origin and Destination Survey (DB1B) which is available to public for free. Furthermore, the authors also offer the cleaned dataset as supplementary materials online, which makes the data more accessible.

Particularly, Table.**??** gives the number of incumbents in each market, and those with less than 2 takes 37.93%.

---

[5]The reason that we rewrite Eq.(**??**) in the form of conditional expectation is for constructing sample analog, for which the key step is calculating the conditional expectation of ϵ_{im}. When there are multiple equilibria, that conditional expectation depends on the equilibrium selection mechanism, and that is why we include $\mathcal{S}$ in the condition.

No. of incumbents	0	1	2	3	4	5	6
	200	840	711	431	327	205	28

7 Example

$$\pi_i = \alpha - \beta y_{-i} - c_i, \quad \text{where } c_i \sim U[0,1]$$

$$\begin{cases} E[\alpha - \beta - c_i | y_i = 1, y_{-i} = 1] > 0 \\ E[\alpha - c_i | y_i = 1, y_{-i} = 0] > 0 \\ E[\alpha - \beta - c_i | y_i = 0, y_{-i} = 1] < 0 \\ E[\alpha - c_i | y_i = 0, y_{-i} = 0] < 0 \end{cases}$$

$$\Longleftrightarrow$$

$$\begin{cases} (\alpha - \beta)Prob\{y_i = 1, y_{-i} = 1\} - \int_{G_{1,1}} c_i f(c_i, c_{-i}) dc_i dc_{-i} > 0 \\ \alpha Prob\{y_i = 1, y_{-i} = 0\} - \int_{G_{1,0}} c_i f(c_i, c_{-i}) dc_i dc_{-i} > 0 \\ -(\alpha - \beta)Prob\{y_i = 0, y_{-i} = 1\} + \int_{G_{0,1}} c_i f(c_i, c_{-i}) dc_i dc_{-i} > 0 \\ -\alpha Prob\{y_i = 0, y_{-i} = 0\} + \int_{G_{0,0}} c_i f(c_i, c_{-i}) dc_i dc_{-i} > 0 \end{cases}$$

where $G_{y_i, y_{-i}} = \{c_i : Y_i = y_i \,\&\, Y_{-i} = y_{-i}\}$

Taking $E[\alpha - c_i | y_i = 1, y_{-i} = 0] > 0 \Longleftrightarrow \alpha Prob\{y_i = 1, y_{-i} = 0\} - \int_{G_{1,0}} c_i f(c_i) dc_i > 0$ as an example,

$$\begin{aligned} & E[\alpha - c_i | y_i = 1, y_{-i} = 0] > 0 \\ \Longleftrightarrow \quad & E[\alpha - c_i | y_i = 1, y_{-i} = 0] Prob\{y_i = 1, y_{-i} = 0\} > 0 \end{aligned} \tag{18}$$

The equivalence is implied by $Prob\{y_i = 1, y_{-i} = 0\} > 0$. We then rewrite Eq.(**??**) as

$$\begin{aligned} & \alpha Prob\{y_i = 1, y_{-i} = 0\} - E[c_i | y_i = 1, y_{-i} = 0] Prob\{y_i = 1, y_{-i} = 0\} \\ = \quad & \alpha Prob\{y_i = 1, y_{-i} = 0\} - \frac{\int_{G_{1,0}} c_i f(c_i, c_{-i}) dc_i dc_{-i}}{Prob\{y_i = 1, y_{-i} = 0\}} Prob\{y_i = 1, y_{-i} = 0\} \\ = \quad & \alpha Prob\{y_i = 1, y_{-i} = 0\} - \int_{G_{1,0}} c_i f(c_i, c_{-i}) dc_i dc_{-i} \end{aligned}$$

Specifically, given the iid uniform distribution assumption on c, $f(c_i, c_{-i}) = 1$. Furthermore, due to multiple equilibria, $G_{1,0}$ is unknown, which makes the calculation of $\int_{G_{1,0}} c_i f(c_i) dc_i dc_{-i}$ infeasible. We will substitute $G_{1,0}$ with its sufficient condition, i.e., the area where $y = (1, 0)$ is the unique equilibrium.

In this specific case, there are two scenarios when $y = (1, 0)$ is the unique solution: first, firm i enters regardless of firm j's action and firm j doesn' enter; second, firm j doesn't enter regardless firm i's action, and firm i enters only when firm j doesn't. Mathematically,

$$\begin{aligned} \int_{G_{1,0}} c_i dc_i dc_{-i} \quad > \quad & \int_{\alpha-\beta}^{1} \int_{0}^{\alpha-\beta} c_i dc_i dc_{-i} + \int_{\alpha}^{1} \int_{\alpha-\beta}^{\alpha} c_i dc_i dc_{-i} \\ = \quad & (1 - (\alpha - \beta)) \frac{(\alpha - \beta)^2}{2} + (1 - \alpha) \frac{(2\alpha - \beta)\beta}{2} \end{aligned}$$

Therefore, the system of inequalities can be rewritten as

$$\begin{cases} (\alpha - \beta) P_{11} - (\alpha - \beta) \frac{(\alpha - \beta)^2}{2} > 0 \\ \alpha P_{10} - \{(1 - (\alpha - \beta)) \frac{(\alpha - \beta)^2}{2} + (1 - \alpha) \frac{(2\alpha - \beta)\beta}{2}\} > 0 \\ -(\alpha - \beta) P_{01} + \{\alpha \frac{1 - \alpha^2}{2} + (\alpha - \beta) \frac{(2\alpha - \beta)\beta}{2}\} > 0 \\ -\alpha P_{00} + (1 - \alpha) \frac{1 - \alpha^2}{2} > 0 \end{cases}$$

Observation should be at market level.

Objective function

$$\begin{aligned}
Q(\theta) \quad = \quad & \|((\alpha - \beta)P_{11} - \frac{(\alpha - \beta)^2}{2}(\alpha - \beta))_-\|+ \\
& \|(\alpha P_{10} - \{(1 - (\alpha - \beta))\frac{(\alpha - \beta)^2}{2} + (1 - \alpha)\frac{(2\alpha - \beta)\beta}{2}\})_-\|+ \\
& \|((\alpha - \beta)P_{0,1} - \{(1 - (\alpha - \beta))\frac{(\alpha - \beta)^2}{2} + (1 - \alpha)\frac{(2\alpha - \beta)\beta}{2}\})_+\|+ \\
& \|(\alpha P_{0,0} - (1 - \alpha)\frac{1 - \alpha^2}{2})_+\|
\end{aligned}$$

Trying to incorporate the information that $E[c_i] = \frac{1}{2}$[6]

$$\begin{aligned}
E[c_i] \quad &= \quad E\Big[E[c_i|y_{-i} = 1]Prob(y_i = 1) + E[c_i|y_{-i} = 0]Prob(y_i = 0)\Big] \\
&= \quad E\Big[E[c_i|y_{-i} = 1]E[y_{-i}] + E[c_i|y_{-i} = 0]E[1 - y_{-i}]\Big] \\
&= \quad E\Big[y_{-i}E[c_i|y_{-i} = 1] + (1 - y_{-i})E[c_i|y_{-i} = 0]\Big] \\
&= \quad \frac{1}{2}
\end{aligned} \tag{19}$$

Furthermore,

$$\begin{aligned}
E[c_i|y_{-i}] \quad &= \quad E[E[c_i|y_i, y_{-i}]|y_{-i}] \\
&= \quad E[E[c_i|y_i = 1, y_{-i}]Prob(y_i = 1|y_{-i}) + E[c_i|y_i = 0, y_{-i}]Prob(y_i = 0|y_{-i})|y_{-i}] \\
&= \quad E[E[c_i|y_i = 1, y_{-i}]E[y_i|y_{-i}] + E[c_i|y_i = 0, y_{-i}]E[1 - y_i|y_{-i}]|y_{-i}] \\
&= \quad E[y_iE[c_i|y_i = 1, y_{-i}] + (1 - y_i)E[c_i|y_i = 0, y_{-i}]|y_{-i}]
\end{aligned} \tag{20}$$

Given $E[\alpha - \beta y_{-i} - c_i|y_i = 1, y_{-i}] > 0$,

$$\begin{aligned}
& E[y_i(\alpha - \beta y_{-i} - c_i)|y_{-i}] > 0 \\
\implies \quad & E[y_i(\alpha - \beta y_{-i}) - y_iE[c_i|y_i = 1, y_{-i}]|y_{-i}] > 0 \\
\implies \quad & E[y_i(\alpha - \beta y_{-i}) - c_i + (1 - y_i)E[c_i|y_i = 0, y_{-i}]|y_{-i}] > 0, \qquad \text{by Eq.(\textbf{??})} \\
\implies \quad & E\Big[y_{-i}E\Big[y_i(\alpha - \beta y_{-i}) - c_i + (1 - y_i)E[c_i|y_i = 0, y_{-i}]\Big|y_{-i} = 1\Big]\Big] \\
= \quad & E\Big[y_{-i}\Big\{E[y_i(\alpha - \beta)|y_{-i} = 1] - E[c_i|y_{-i} = 1] + E[1 - y_i|y_{-i} = 1]E[c_i|y_i = 0, y_{-i} = 1]\Big\}\Big] > 0
\end{aligned} \tag{21}$$

Similarly,

$$E\Big[(1 - y_{-i})\Big\{E[y_i\alpha|y_{-i} = 0] - E[c_i|y_{-i} = 0] + E[1 - y_i|y_{-i} = 0]E[c_i|y_i = 0, y_{-i} = 0]\Big\}\Big] > 0 \tag{22}$$

Combining Eq.(**??**) and Eq.(**??**), by Eq.(**??**), we have

$$\begin{aligned}
E\Big[& y_{-i}\Big\{(\alpha - \beta)E[y_i|y_{-i} = 1] + E[1 - y_i|y_{-i} = 1]]E[c_i|y_i = 0, y_{-i} = 1]\Big\} \\
& + (1 - y_{-i})\Big\{\alpha E[y_i|y_{-i} = 0] + E[1 - y_i|y_{-i} = 0]E[c_i|y_i = 0, y_{-i} = 0]\Big\}\Big] - \frac{1}{2} > 0
\end{aligned} \tag{23}$$

Substitute the expected entry decisions with the observed probabilities,

$$(\alpha - \beta)P_{11} + P_{01}E[c_i|y_i = 0, y_{-i} = 1] + \alpha P_{10} + P_{00}E[c_i|y_i = 0, y_{-i} = 0] - \frac{1}{2} > 0 \tag{24}$$

Given that $P_{01}E[c_i|y_i = 0, y_{-i} = 1] < \alpha\frac{1 - \alpha^2}{2} + (\alpha - \beta)\frac{(2\alpha - \beta)\beta}{2}$ and $E[c_i|y_i = 0, y_{-i} = 0] = \frac{1 + \alpha}{2}$,

$$(\alpha - \beta)P_{11} + \alpha\frac{1 - \alpha^2}{2} + (\alpha - \beta)\frac{(2\alpha - \beta)\beta}{2} + \alpha P_{10} + \frac{1 + \alpha}{2}P_{00} - \frac{1}{2} > 0 \tag{25}$$

Similarly, given $E[\alpha - \beta y_{-i} - c_i|y_i = 0, y_{-i}] < 0$,

---

[6]Motivation: trying to exclude the case where $\alpha = \beta = 0$. If that is the case, we should observe $P_{00} = 1$ and $P_{11} = P_{10} = P_{01} = 0$. In terms of conditional expectation, it leads to $E[c_i] = E[c_i|y_i = 0, y_{-i} = 0]P_{00} + ... = \frac{1 + \alpha}{2}P_{00} + ... = \frac{1}{2}$, given $\alpha = \beta = 0$, $\frac{1}{2}P_{00} = \frac{1}{2}$, which would be violated unless $P_{00} = 1$

$$E\left[y_{-i}\left\{E[(1-y_i)(\alpha-\beta)|y_{-i}=1] - E[c_i|y_{-i}=1] + E[y_i|y_{-i}=1]E[c_i|y_i=1,y_{-i}=1]\right\}\right] < 0 \qquad (26)$$

and

$$E\left[(1-y_{-i})\left\{E[(1-y_i)\alpha|y_{-i}=0] - E[c_i|y_{-i}=0] + E[y_i|y_{-i}=0]E[c_i|y_i=1,y_{-i}=0]\right\}\right] < 0 \qquad (27)$$

Combining Eq.(**??**) and Eq.(**??**),

$$E\left[y_{-i}\left\{(\alpha-\beta)E[1-y_i|y_{-i}=1] + E[y_i|y_{-i}=1]E[c_i|y_i=1,y_{-i}=1]\right\}\right.$$
$$\left.(1-y_{-i})\left\{\alpha E[1-y_i|y_{-i}=0] + E[y_i|y_{-i}=0]E[c_i|y_i=1,y_{-i}=0]\right\}\right] < 0 \qquad (28)$$

Substitute the expected entry decisions with the observed probabilities,

$$(\alpha-\beta)P_{01} + P_{11}E[c_i|y_i=1,y_{-i}=1] + \alpha P_{00} + P_{10}E[c_i|y_i=1,y_{-i}=0] - \frac{1}{2} < 0 \qquad (29)$$

Given that $E[c_i|y_i=1,y_{-i}=1] = \dfrac{\alpha-\beta}{2}$ and $P_{10}E[c_i|y_i=1,y_{-i}=0] < (1-(\alpha-\beta))\dfrac{(\alpha-\beta)^2}{2} + (1-\alpha)\dfrac{(2\alpha-\beta)\beta}{2}$,

$$(\alpha-\beta)P_{01} + \frac{\alpha-\beta}{2}P_{11} + \alpha P_{00} + (1-(\alpha-\beta))\frac{(\alpha-\beta)^2}{2} + (1-\alpha)\frac{(2\alpha-\beta)\beta}{2} - \frac{1}{2} < 0 \qquad (30)$$

Information:

1. Profit Inequalities

2. $E[c_i] = \dfrac{1}{2}$

3. $P_{00} = (1-\alpha)^2$, $P_{11} = (\alpha-\beta)^2$

Questions:

1. How to incorporate all inequalities?

2. Any information reused?

3. How to determine between observed probabilities and predicted ones?

4. Conditional (profit inequalities) v.s. unconditional moments ($E[c_i] = \dfrac{1}{2}$)

Consider the case

$$\pi_i = \alpha x_i - \beta y_{-i} - c_i, \quad \text{where } c_i \sim U[0,1]$$

$$\begin{aligned}
& E[\alpha x_i - \beta y_{-i} - c_i|y_i=1,y_{-i}=1,x] > 0 \\
\iff\ & E[y_i y_{-i}(\alpha x_i - \beta y_{-i} - c_i)|x] > 0 \\
\iff\ & E\left[y_i y_{-i}E[(\alpha x_i - \beta y_{-i} - c_i)|y_i=1,y_{-i}=1,x]\big|x\right] > 0 \\
\iff\ & E\left[y_i y_{-i}\{(\alpha x_i - \beta) - E[c_i|y_i=1,y_{-i}=1,x]\}\big|x\right] > 0 \\
\iff\ & (\alpha x_i - \beta)E[y_i y_{-i}|x] - E[c_i|y_i=1,y_{-i}=1,x]E[y_i y_{-i}|x] > 0 \\
\iff\ & (\alpha x_i - \beta)Prob\{y_i=1,y_{-i}=1|x\} - \int_0^{\alpha x_{-i}-\beta}\int_0^{\alpha x_i - \beta} c_i f(c_i,c_{-i})dc_i dc_{-i} > 0 \\
\iff\ & (\alpha x_i - \beta)P_{11|x} - (\alpha x_{-i} - \beta)\frac{(\alpha x_i - \beta)^2}{2} > 0
\end{aligned} \qquad (31)$$

Following a similar process, we will reach to

$$
\begin{cases}
(\alpha x_1 - \beta)P_{11|x} - (\alpha x_2 - \beta)\dfrac{(\alpha x_1 - \beta)^2}{2} > 0 \\[2mm]
\alpha x_1 P_{10|x} - \{(1 - (\alpha x_2 - \beta))\dfrac{(\alpha x_1 - \beta)^2}{2} + (1 - \alpha x_2)\dfrac{(2\alpha x_1 - \beta)\beta}{2}\} > 0 \\[2mm]
-(\alpha x_1 - \beta)P_{01|x} + \{\alpha x_2 \dfrac{1 - (\alpha x_1)^2}{2} + (\alpha x_2 - \beta)\dfrac{(2\alpha x_1 - \beta)\beta}{2}\} > 0 \\[2mm]
-\alpha x_1 P_{00|x} + (1 - \alpha x_2)\dfrac{1 - (\alpha x_1)^2}{2} > 0
\end{cases}
$$

$$
\begin{cases}
(\alpha x_2 - \beta)P_{11|x} - (\alpha x_1 - \beta)\dfrac{(\alpha x_2 - \beta)^2}{2} > 0 \\[2mm]
-(\alpha x_2 - \beta)P_{10|x} + \{\alpha x_2 \dfrac{1 - (\alpha x_2)^2}{2} + (\alpha x_1 - \beta)\dfrac{(2\alpha x_2 - \beta)\beta}{2}\} > 0 \\[2mm]
\alpha x_2 P_{01|x} - \{(1 - (\alpha x_1 - \beta))\dfrac{(\alpha x_2 - \beta)^2}{2} + (1 - \alpha x_1)\dfrac{(2\alpha x_2 - \beta)\beta}{2}\} > 0 \\[2mm]
-\alpha x_2 P_{00|x} + (1 - \alpha x_1)\dfrac{1 - (\alpha x_2)^2}{2} > 0
\end{cases}
$$

And the counterpart for Eq.(**??**) 7 Eq.(**??**) is

$$
(\alpha x_1 - \beta)P_{11|x} + \alpha x_2 \frac{1 - (\alpha x_1)^2}{2} + (\alpha x_2 - \beta)\frac{(2\alpha x_1 - \beta)\beta}{2} + \alpha x_1 P_{10|x} + \frac{1 + \alpha x_1}{2}P_{00|x} - \frac{1}{2} \quad > 0
$$

$$
(\alpha x_1 - \beta)P_{01|x} + \frac{\alpha x_1 - \beta}{2}P_{11|x} + \alpha x_1 P_{00|x} + (1 - (\alpha x_2 - \beta))\frac{(\alpha x_1 - \beta)^2}{2} + (1 - \alpha x_2)\frac{(2\alpha x_1 - \beta)\beta}{2} - \frac{1}{2} \quad < 0
$$

$$\tag{32}$$