

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

Shuang Wang

March 11, 2018

1 Introduction

In various empirical problems, such as entry games, discrete choice model with social interactions and network formation, the observed outcome could be any one of the 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. However, econometricians know little about how a certain outcome is selected in reality. Therefore, 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).

Consequently, moment inequality emerges as a method to estimate such incomplete models. According to Pakes [2010], there are two different ways to construct inequalities and they are based on different sets of assumptions.

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 interacting games. The ideas behind this approach date to Tamer [2003], and are developed in more details in papers by Ciliberto and Tamer [2009] and Andrews et al. [2004]. The basic idea of GDC is constructing upper and lower bounds for the probability of each possible equilibrium outcome. Since it is generally impossible to analytically derive the bounds, they usually use simulation, for which distributional assumption on the unobservables is inevitable. For example, in Ciliberto and Tamer [2009], they assume that the unobserved part in airlines' profit function is consistent of four component, and each of them is independent and normally distributed.¹ Since econometricians do not have much theory about the unobservables, such a restrictive distributional assumption can possibly induce misleading estimates.

The second approach is based on the inequalities generated by the assumption that expected profits from the observed choice must be at least as great as those from alternative feasible choices, so it is referred as the *profit inequality* approach (PI henceforth). The approach considered here is a direct extension of the work in Pakes et al. [2015]. Unlike GDC, PI does not require distributional assumptions on the unobservables, although it requires the existence of a weight function that helps us circumvent the unobservables without violating the inequalities, or there would be *selection bias* problem. More details on the common assumptions shared by GDC and PI as well as their distinguishing assumptions can be found in Section 3.

Therefore, the objective of this paper is using the method of PI to reestimate the model in Ciliberto and Tamer [2009] and checking whether the estimates are robust to changes in the distributional assumptions. Also, 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 which is induced by agent-specific structural errors. Explicitly, although PI does not require distribution assumption on the unobservable part of the profit function, it does require a weight function that makes

¹In Section 5.2 of Ciliberto and Tamer [2009], the independent assumption is relaxed, but the normal assumption is retained throughout the paper.

the weighted sum of the unobservables negligible in the inequalities. Therefore, if the unobservables are centered at zero in expectation, its sample analogue will converge to zero as desired. However, since part of the "unobservables" is actually observable to the agents and affects their decisions, we only have a biased sample of them. So this part of the "unobservables" is usually referred as *structural error* in literature. For instance, in the entry game among airlines, it is likely that the firms entered the market because of an unobservable firm-level positive demand shock. If we researchers only have data on the entrants, the sample average of the unobservables over the observations may converges to a positive number, which could possibly turn the inequality over if ignored.

In the existing literature, there are generally two approaches addresses this *selection bias* problem. The first one is assuming that the unobservables is some market-level fixed effects, so we can difference the unobservables out by differencing the inequalities for two agents in the same market. (Ellickson et al. [2013], Ho and Pakes [2014]). The second approach is assuming that the unobservables do not vary with agents' decisions. For example, in Pakes et al. [2015], they use the example of ATM network in Ishii [2008] to illustrate the method of PI. In that model, the marginal installation cost of each bank has an idiosyncratic deviation from the average level, and it is unobservable to researchers. However, since it does not vary with the number of ATM installed by the bank, its sample analog across all banks in the market goes to zero in probability. Similar approach is adopted in Ho [2009]. However, since in Ciliberto and Tamer [2009], the structural error is firm-and-market specific, none of the two approaches mentioned above is feasible.

As far as I know, the only work that allows for agent-level, decision-dependent unobservables is Eizenberg [2014]. He estimated 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. His method is inapplicable for the entry games discussed here because (1) he assumes that the cost for unlisted products is bounded above; (2) the characteristics of unlisted products are known. They are probably plausible in that context because potential products considered there are CPU selected from a finite set of options installed in existing notebook PC product lines. However, characteristics such as *airport presence* can be unobservable for potential entrants in airline markets and the decision of not entering might have been characterized by an infinitely-large negative demand shock.

Therefore, we need a new approach to solve the selection bias problem which is applicable in the context of airline markets or potentially other empirical applications.

Second, to my knowledge, there is no existing works that compare the results from GDC and PI with real data. The most closely-related work is Pakes [2010]. The author listed different assumptions of GDC and PI, and did a Monte Carlo experiment to check their robustness. 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 Ciliberto and Tamer [2009] 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 Ciliberto and Tamer [2009], I will follow their basic settings and introduce the implication of GDC and PI assumptions in this specific case.

The main idea of Ciliberto and Tamer [2009] is estimating the heterogeneity competitive effects among airlines, i.e., how would one firm's entry affects the profit of other firms. Specifically, suppose there are K potential entrants indexed by $i = 1, \dots, K$ 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. 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}).

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

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

3.1 Common Assumptions

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

Assumption 1: We assume

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

where function $y(\cdot)$ gives airline i 's strategy given its information set \mathcal{J}_{im} , \mathbf{y}_{im} is 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} . Therefore, we need a "structural" model characterizing such interaction for future counterfactual analysis, like predicting what expected profits would be if the agent deviates from his observed choice. To simplify the analysis, we assume agent's perception about other agents' strategies and characteristics does not change with its own actions, given the information set 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} ,

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

Then, in accordance with Ciliberto and Tamer [2009], we further assume agents have complete information,

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

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

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 $v_{i,m,y}^1$, where $v_{i,m,y}^1 = \pi(y_{im}, y_{-im}, X_{im}) - E[\pi(y_{im}, \mathbf{y}_{-im}, \mathbf{X}_{im}) | \mathcal{J}_{im}]$. Here, we trade such generality for comparability.

Assumption 3 implies that Eq.(2) can be further simplified as

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

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 $r(y_{im}, y_{-im}, X_{im}^o; \theta)$. 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_{im}^o) and is identified up to a finite set of parameters θ . Following Pakes et al. [2015], we define such structural error as

$$v_{i,m,y}^2 = \pi(y_{im}, y_{-im}, X_{im}) - r(y_{im}, y_{-im}, X_{im}^o, \theta). \quad (4)$$

²Throughout 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 (Pakes [2010]) and additional assumptions are needed for estimation. Here, for simplicity, we assume "correct expectation", following Pakes et al. [2015], i.e., $E[\cdot] = \mathcal{E}[\cdot]$

Therefore, in Eq.(1),

$$r(y_{im}, y_{-im}, X_{im}^o, \theta) = y_{im}(S'_m \alpha_i + Z'_{im} \beta_i + W'_{im} \gamma_i + \sum_{j \neq i} \delta_j^i y_{jm} + \sum_{j \neq i} Z'_{jm} \phi_j^i y_{jm}) \quad (5)$$

$$v_{i,m,y}^2 = y_{im} \epsilon_{im} \quad (6)$$

In [Ciliberto and Tamer \[2009\]](#), in order to compute 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

$$\epsilon_{im} = u_m + u_m^o + u_m^d + u_{im},$$

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

With such an assumption, [Ciliberto and Tamer \[2009\]](#) use simulation to calculate the equilibria where the condition in Eq. (3) is satisfied. Explicitly, for each draw of $\{\epsilon_{im}\}_{i,m}$, if $y_m = \{y_{im}, y_{-im}\}$ is the unique equilibrium, we add one to both the upper and lower bounds for the probability of y_m ; if it is one of the multiple equilibria, we add one to the upper bound only, then we take the average over all random draws. Then those parameter values that make the inequalities hold are in the parameter set.

Rather than constructing bounds for outcome probabilities, [Pakes et al. \[2015\]](#) directly uses Eq.(3) for estimation. Since for any airline i , the profit is zero if it chooses not to enter the market, i.e., $\pi(0, y_{-im}, X_{im}) = 0$, we have

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

which is equivalent to

$$(-1)^{y_{im}+1} [r(1, y_{-im}, X_{im}^o, \theta) + \epsilon_{im}] \geq 0 \quad (7)$$

However, ϵ_{im} is unobservable to researchers. To circumvent it in estimation, [Pakes et al. \[2015\]](#) makes the following assumption.

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

$$E[\sum_i h(y_{im}, y_{-im}, X_{im}^o) (-1)^{y_{im}+1} \epsilon_{im}] \leq 0 \quad (8)$$

Assumption PI implies that if we sum Eq.(7) over airlines using $h(y_{im}, y_{-im}, X_{im}^o)$ as a weight function and then take expectation, we will have

$$E[\sum_i h(y_{im}, y_{-im}, X_{im}^o) (-1)^{y_{im}+1} r(1, y_{-im}, X_{im}^o, \theta)] \geq 0 \quad (9)$$

Next, with observations of different markets, we can replace the expectation with sample analogue to estimate θ .

4 Issues in applying PI

4.1 Selection Bias

However, Assumption PI is infeasible in this specific context due to the *selection bias* problem. Since the structural error ϵ_{im} here is firm-and-market specific and it is observable to the airline, its expectation varies with the firm's entry decision. Specifically, define I_m as the set of incumbents in market m , i.e., $I_m = \{k : y_{km} = 1\}$. Then Eq.(8) in Assumption PI can be rewritten as

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

Also,

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

Similarly,

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

Then, given $E[\epsilon_{im}] = 0$, $h(y_{im}, y_{-im}, X_{im}^o) \geq 0$, we would expect Eq.(10) to be positive, which violates Assumption PI. Since ϵ_{im} is firm-and-market-specific, I cannot eliminate it by differencing two inequalities with the same ϵ as previous literature did. Therefore, I need a new approach to solve this selection bias problem.

$E[\epsilon_{im}] = 0$ v.s. ϵ is mean-independent to the observables

4.2 Two-stage Estimation

The empirical strategy proposed by Pakes et al. [2015] is applicable only to estimate a scalar parameter in a linear model. Specifically, they complete the estimation by two steps. First, they achieved point identification for nuisance parameters using a structural demand estimation model. Then, they constructed upper and lower bounds with the estimated nuisance parameter plugged in for the left scalar parameter which entered linearly in the model³.

In Ciliberto and Tamer [2009], the parameter set θ includes $\{\alpha_i, \beta_i, \gamma_i, \{\delta_j^i\}_j, \{\phi_j^i\}_j\}_i$. Since they only have a reduced form profit function, it is infeasible to do a two-stage estimation here. Therefore, I proposed a pseudo two-stage estimation procedure described in Section.5

5 Preliminary Empirical Strategy

For now, let's focus on 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 $\forall i, j$. Then

$$\pi(y_{im}, y_{-im}, X_{im}) = y_{im}(S'_m \alpha + Z'_{im} \beta + W'_{im} \gamma + \sum_{j \neq i} \delta_j y_{jm}) + y_{im} \epsilon_{im} \quad (10)$$

In the first stage, we can utilize the markets with no firms or only one firm to get point identification for the nuisance parameters under the framework of a semi-parametric binary choice model⁴.

First-stage For all firms in the markets with no entry and the entering firm in the markets with only one entrant,

$$\begin{aligned}
y_{im} &= 1[S'_m \alpha + Z'_{im} \beta + W'_{im} \gamma + \epsilon_{im} \geq 0] \\
\Rightarrow E[y_{im} | X_{im}^o] &= \text{Prob}[S'_m \alpha + Z'_{im} \beta + W'_{im} \gamma + \epsilon_{im} \geq 0] \\
&= \text{Prob}[\epsilon_{im} \geq -S'_m \alpha - Z'_{im} \beta - W'_{im} \gamma] \\
&= 1 - F(-S'_m \alpha - Z'_{im} \beta - W'_{im} \gamma)
\end{aligned}$$

Then

³In the earlier version of Pakes et al. [2015], they also proposed an approach to partially identify a vector of parameters. That is another strand that I will check in the future.

⁴In Ciliberto and Tamer [2009], the components in Z_{im} and W_{im} are constructed at endpoint level data so that they are still "observable" even if the airline is not in the market. For instance, for the *airline presence* in Z_{im} , they first compute a carrier's ratio of markets served by an airline out of an airport over the total number of markets served out of an airport by at least one carrier. Then, they define the carrier's airport presence as the average of the carrier's airport presence at the two endpoints.

$$y_{im} = E[y_{im}|X_{im}^o] + u_{im} \quad \text{and} \quad E[u_{im}|X_{im}^o] = 0 \quad \text{by definition}$$

This suggests a least square estimator. Furthermore, without distributional assumption on ϵ_{im} , $E[y_{im}|X_{im}^o]$ is unknown. Following Ichimura [1993], I am going to replace the $F(-S'_m\alpha - Z'_{im}\beta - W'_{im}\gamma)$ in it with a kernel estimator, which gives the estimators a semiparametric nature.

Specifically,

$$\hat{\psi} = \arg \max_{\psi} \sum_{i=1}^N \sum_{m=1}^M \mathbb{1}(X_{im} \in \mathbf{X}) (y_{im} - 1 + \hat{F}(-S'_m\alpha - Z'_{im}\beta - W'_{im}\gamma))^2$$

where

- $\psi = [\alpha, \beta, \gamma]$,
- $\hat{F}(-S'_m\alpha - Z'_{im}\beta - W'_{im}\gamma)$ is the kernel estimator
- $\mathbb{1}(X_{im} \in \mathbf{X})$ is a trimming function which is introduced to guarantee that the density of $(-S'_m\alpha - Z'_{im}\beta - W'_{im}\gamma)$ is bounded away from 0 on \mathbf{X} .

Second-stage Then, we plug $\hat{\psi}$ into the profit function, and then construct the feasible set for each δ_j following the method in Pakes et al. [2015]. Take upper bound as an example, consider market m and market m' , where $I_m = I_{m'}$

$$\pi_{im} = S'_m\hat{\alpha} + Z'_{im}\hat{\beta} + W'_{im}\hat{\gamma} + \sum_{k \neq i, k \in I_m} \delta_k + \epsilon_{im} \geq 0, \quad \forall i \in I_m \quad (11)$$

$$\pi_{jm'} = S'_{m'}\hat{\alpha} + Z'_{jm'}\hat{\beta} + W'_{jm'}\hat{\gamma} + \sum_{k \in I_{m'}} \delta_k + \epsilon_{jm'} \leq 0, \quad \forall j \notin I_{m'} \quad (12)$$

Subtract Eq.(11) from Eq.(12), we will have

$$\delta_i \leq (S'_{m'} - S'_m)\hat{\alpha} + (Z'_{im'} - Z'_{im})\hat{\beta} + (W'_{im'} - W'_{im})\hat{\gamma} + \epsilon_{jm'} - \epsilon_{im} \quad (13)$$

Take expectation for both sides of Eq.(13),

$$\delta_i \leq (S'_{m'} - S'_m)\hat{\alpha} + (Z'_{im'} - Z'_{im})\hat{\beta} + (W'_{im'} - W'_{im})\hat{\gamma} + E[\epsilon_{im}|\pi_{im} > 0] - E[\epsilon_{jm'}|\pi_{jm'} < 0] \quad (14)$$

where

$$\begin{aligned} E[\epsilon_{im}|\pi_{im} > 0] &= E[\epsilon_{im}|\epsilon_{im} \geq -(S'_m\hat{\alpha} + Z'_{im}\hat{\beta} + W'_{im}\hat{\gamma} + \sum_{k \neq i, k \in I_m} \delta_k)] \\ E[\epsilon_{jm'}|\pi_{jm'} < 0] &= E[\epsilon_{jm'}|\epsilon_{jm'} \leq -(S'_{m'}\hat{\alpha} + Z'_{jm'}\hat{\beta} + W'_{jm'}\hat{\gamma} + \sum_{k \in I_{m'}} \delta_k)] \end{aligned}$$

Using the estimated density $\hat{F}(\cdot)$, $E[\epsilon_{im}|\pi_{im} > 0]$ and $E[\epsilon_{jm'}|\pi_{jm'} < 0]$ are identified up to the competitive effects, $\{\delta_i\}_i$. Therefore, the parameter set should include any values $\{\delta_i\}_i$ conforming to the inequality above.

6 Data

In Ciliberto and Tamer [2009], 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.

7 The easiest PI estimation:

As discussed in Section 4, the usual way in which PI deals with structural error $v_{i,m,y}^2$ is differencing across either markets or firms. Therefore, when the error is both market- and firm- specific so that differencing is infeasible, there could be selection problem. For now, we circumvent this complexity by assuming there is no $v_{i,m,y}^2$, instead, we allow for a measurement error in $cost$, which is an element of W'_{im} in [Ciliberto and Tamer \[2009\]](#).

$$\begin{aligned}
\pi(y_{im}, y_{-im}, X_{im}) &= y_{im}(S'_m \alpha_i + Z'_{im} \beta_i + W'_{im} \gamma_i + \sum_{j \neq i} \delta_j^i y_{jm} + \sum_{j \neq i} Z'_{jm} \phi_j^i y_{jm}) \\
&= y_{im}(S'_m \alpha_i + Z'_{im} \beta_i + \sum_{j \neq i} \delta_j^i y_{jm} + \sum_{j \neq i} Z'_{jm} \phi_j^i y_{jm}) + y_{im} W'_{im} \gamma_i \\
&= y_{im}(S'_m \alpha_i + Z'_{im} \beta_i + \sum_{j \neq i} \delta_j^i y_{jm} + \sum_{j \neq i} Z'_{jm} \phi_j^i y_{jm}) + (y_{im} W'_{im} \gamma_i + \varphi_{im}) \gamma_i \\
&= y_{im}(S'_m \alpha_i + Z'_{im} \beta_i + W'_{im} \gamma_i + \sum_{j \neq i} \delta_j^i y_{jm} + \sum_{j \neq i} Z'_{jm} \phi_j^i y_{jm}) + \varphi_{im} \gamma_i \\
&= r(y_{im}, y_{-im}, X_{im}^o, \theta) + v_{i,m}^1
\end{aligned} \tag{15}$$

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

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

$$r(y_{im}, y_{-im}, X_{im}^o, \theta) = y_{im}(S'_m \alpha + Z'_{im} \beta + W'_{im} \gamma + \sum_{j \neq i} \delta_j y_{jm})$$

Then, Eq.(7) can be rewritten as

$$(-1)^{y_{im}+1} (r(1, y_{-im}, X_{im}^o, \theta) + v_{i,m}^1) \geq 0 \tag{16}$$

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

$$E[(-1)^{y_{im}+1} r(1, y_{-im}, X_{im}^o, \theta) | z_{im}] \geq 0 \tag{17}$$

for any $z_{im} \in J_{im}$. Translating expectations into sample means, the equations for estimation is therefore

$$\frac{1}{MI} \sum_m \sum_i (-1)^{y_{im}+1} r(1, y_{-im}, X_{im}^o, \theta) \otimes g(z_{im}) \geq 0 \tag{18}$$

And the objective function is

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

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

Then we can construct the confidence sets for θ following ? as [Ciliberto and Tamer \[2009\]](#) does.

To do

1. Check the dataset
2. Roadmap PI may allow for higher dimension of heterogeneity. It would be interesting to see that the results are different under different assumptions. Furthermore, it would be even more interesting if there is heterogeneity among LCC airlines.
3. Why or whether PI and GDC give different confidence sets?

With distributional assumption, GDC gives a narrower confidence sets. Intuitively, it is similar to MLE achieves Cramer-Rao lower bound when the sample size tends to infinity. This means that no consistent estimator has lower asymptotic mean squared error than the MLE (or other estimators attaining this bound).

Perhaps we could start with binary probit model where there is also selection bias problem ? Is it possible to estimate it without distributional assumption?

consider Andrews and Shi [2013]

4. Whether PI is less computationally-burdensome?

For PI, it seems that we only need to calculate $2 \times k$ profits; however, for GDC, if we only calculate the upper and lower bounds for the observed outcome in each market, $k \times R$ profits are needed. Furthermore, to calculate the observed probability, for each outcome, a frequency estimator is needed for each $4 \times \text{no. of continuous explanatory variables}$.

why moment inequality

trade-off

- Ciliberto and Tamer [2009]: multi-agent game with multiple equilibria with no restrictions on player heterogeneity and equilibria selection mechanism
 - without making equilibrium selection assumptions
 - static, complete information
 - mixed strategy (auction) v.s. multiple equilibrium
 - reduced-form profit function
 - assumption: when firm i is to make the entry decision, the market characteristics and the entry decisions of other firms are taken as given. Is it possible to be generalized to a structural, general equilibrium model? also related to counter-factual analysis

References

- Donald WK Andrews and Xiaoxia Shi. Inference based on conditional moment inequalities. *Econometrica*, 81(2):609–666, 2013.
- Donald WK Andrews, Steven Berry, and Panle Jia. Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store location. *manuscript, Yale University*, 2004.
- Victor Chernozhukov, Han Hong, and Elie Tamer. Estimation and confidence regions for parameter sets in econometric models. *Econometrica*, 75(5):1243–1284, 2007.
- Federico Ciliberto and Elie Tamer. Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828, 2009.
- Alon Eizenberg. Upstream innovation and product variety in the us home pc market. *The Review of Economic Studies*, 81(3):1003–1045, 2014.
- Paul B Ellickson, Stephanie Houghton, and Christopher Timmins. Estimating network economies in retail chains: a revealed preference approach. *The RAND Journal of Economics*, 44(2):169–193, 2013.
- Kate Ho and Ariel Pakes. Hospital choices, hospital prices, and financial incentives to physicians. *The American Economic Review*, 104(12):3841–3884, 2014.
- Katherine Ho. Insurer-provider networks in the medical care market. *The American Economic Review*, 99(1):393–430, 2009.
- Hidehiko Ichimura. Semiparametric least squares (sls) and weighted sls estimation of single-index models. *Journal of Econometrics*, 58(1-2):71–120, 1993.
- Joy Ishii. Compatibility, competition, and investment in network industries: Atm networks in the banking industry. *Stanford University WP*, 2008.
- Ariel Pakes. Alternative models for moment inequalities. *Econometrica*, 78(6):1783–1822, 2010.

Ariel Pakes, Jack Porter, Kate Ho, and Joy Ishii. Moment inequalities and their application. *Econometrica*, 83(1):315–334, 2015.

Elie Tamer. Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies*, 70(1):147–165, 2003.