Multiple Equilibria of Aggregated Players in Airline Markets*

Haozhan Gao

November 2023

Abstract

In airline markets, the presence of a hub in either endpoint city can significantly influence an airline's entry decision, as hubs typically offer larger passenger volumes and allow airlines to centralize services and reduce costs through economies of scale. Using pre-Covid data from the first quarter of 2019, I analyze airlines' entries into hub markets, which are defined as markets where at least one endpoint acts as a hub for major airlines. Following the classic framework established by Ciliberto and Tamer (2009), I construct a binary entry game for hub players—which are aggregated from all airlines with a hub at either of the endpoints—and non-hub players comprising the other airlines. Without specifying an equilibrium selection rule, I focus solely on partial identification due to the existence of multiple equilibria. To obtain the confidence intervals of structural parameters in this game, I develop a new estimation strategy by incorporating the moment inequality test method proposed by Cox and Shi (2023), which requires no tuning parameter and is computationally fast. Moreover, the non-convexity of the objective function intrinsically causes the over-rejection issue. Hence, I propose two methods to alleviate this problem by strategically selecting starting parameters and using seeds to control the simulation process, aiming to enhance effectiveness. These methods can be applied to broader scenarios that involve optimizing a non-convex objective function. My results show that compared to non-hub players, hub players' payoffs rely more heavily on market size and market presence and are less prone to the opponent's entry.

1 Introduction

Airlines' entry decisions are often studied within a game structure, where each airline's payoff is a function of market characteristics and the entries of opponents. This is because simple probit/logit regressions of entry decisions suffer from omitted variables or simultaneity issues. Solving the game requires us to formulate entry likelihood functions for the players. However, the existence of multiple equilibria introduces difficulties. By dividing the space for error terms, we can specify the region where multiple equilibria reside. Some literature tackles this problem by specifying an

*I would like to express my sincere gratitude to Prof. Lorenzo Magnolfi and Prof. Xiaoxia Shi for their invaluable advice and guidance on this thesis.

equilibrium selection rule, which usually relies on strong assumptions: Bjorn and Vuong (1984) take an approach of random equilibria selection, Jia (2008) focuses on the most profitable extreme equilibrium, and Berry (1992) allows for heterogeneity among firms and makes assumptions on entry order. Alternatively, partial likelihood can be used in this scenario: Tamer (2003) shows that when the model contains enough information, identification can be achieved 'at the infinity' at rate $\sqrt{n}$. Notably, Ciliberto and Tamer (2009) build a framework that allows for general forms of heterogeneity across players without making equilibrium selection assumptions, focusing only on partial identification. Moreover, Kline and Tamer (2016) develop a Bayesian approach to inference in a class of partially identified models, which are characterized by a known mapping between a point-identified reduced-form parameter and the identified set.

In this paper, I adopt the framework of Ciliberto and Tamer (2009) to study airlines' entry decisions in hub markets. Hub markets are defined as those where at least one endpoint serves as a hub for major airlines. The presence of a hub significantly influences an airline's entry decision due to typically larger passenger volumes and the opportunity for airlines to centralize services and reduce costs through economies of scale. I construct a binary entry game for aggregated players: the hub player, which includes all airlines owning a hub at either market endpoint, and the non-hub player, comprised of other airlines.

Subsequently, I estimate the confidence intervals of structural parameters in their payoff functions using a novel method. Chernozhukov, Hong, and Tamer (2007) offer estimators and confidence regions for parameters that minimize a criterion function $Q(\theta)$. Their method allows a tolerance term and involves subsampling and an iterative process, thus is relatively computation-intensive. In this paper, I propose an alternative way of estimation by incorporating the moment inequality test method proposed by Cox and Shi (2023), which compares the quasi-likelihood ratio statistic to a chi-square critical value and requires no tuning parameter for moment selection, thus making it computationally fast.

Additionally, acknowledging that the non-convexity of the objective function inevitably leads to an over-rejection problem, I seek to mitigate this by strategically selecting starting parameters from the confidence set and using seeds to control the simulation process, aiming to enhance the effectiveness of initial guesses. My estimation results show that this method significantly improve the performance of estimation compared to a random simulation approach.

Furthermore, my results show that compared to non-hub players, hub players' payoffs rely more on market size and market presence and are less sensitive to the entry of opponents.

2 Model

Assume hub- and non-hub player ($i \in \{1, 2\}$) make entry decisions based on the following Bresnahan and Reiss 2×2 game:

$$y_{1m} = \mathbf{1}\{\alpha_1' Z_{1m} + \delta_1 y_{2m} + \epsilon_{1m} > 0\}$$
$$y_{2m} = \mathbf{1}\{\alpha_2' Z_{2m} + \delta_2 y_{1m} + \epsilon_{2m} > 0\}$$

$y_{im} \in \{0, 1\}$ indicates player i's entry decision and they enter only if they earn positive profit. (Z_{1m}, Z_{2m}) is a vector of observed exogenous regressors that contain market-specific characteristics. Specifically, I define $Z_{im} = (s_m, x_m, 1)'$, where s_m is market size defined as the population at the endpoints of the market and x_m is market presence, which is a both market- and player-specific variable: For each market defined by a connection between two endpoint airports, it is obtained by computing the number of markets that the player serves from one airport and divided by the total number of markets served from that airport by any airline, then averaging between two endpoint airports. Notice that x_m is an *excluded regressor* since for each player i, $x_m = x_{im}$ and the payoff of one player is not affected by the other player's market presence. Market presence usually represents a network effect, since an already existed "dense network" will be beneficial to airlines' entry. δ_i capture the competition effect of the opponent's entry and are usually negative. However, when duopoly is more profitable than monopoly, this term can be positive. Further more, I assume $\epsilon_{im} \sim N(0, \sigma^2)$ are i.i.d. with infinite support.

This binary entry game assumes simultaneous entries and thus has multiple equilibria $(1, 0)$ and $(0, 1)$ coexist when $(\epsilon_{1m}, \epsilon_{2m})$ are moderate. Specifically, when $-\alpha_i' Z_i \leq \epsilon_i \leq -\alpha_i' Z_i - \delta_{3-i}$ for $i = \{1, 2\}$, given the other player has entered, the best response is to not enter. Since we are agnostic about player's prior, both equilibria could exist in this region. Figure 1, which is similar to Ciliberto and Tamer (2009), illustrates the equilibria regions for δ being both negative (left) or positive (right).

One caveat is that when δ_1 and δ_2 have different signs, there will be no equilibrium in the multiple equilibria region of the previous cases. Suppose $\delta_1 < 0$ and $\delta_2 > 0$. Given player 2 not entering, player 1 will enter since he will earn a non-negative profit with the absence of competition with player 2. However, given player 1 is entering, player 2 will surely enter as well since player 1's entry make player 2's entry more profitable. We assume away this case for simplicity.

Next, we construct the choice probabilities for 4 possible equilibria $(0,0),(0,1),(1,0)$ and $(1,1)$ based on divisions of the ϵ space:

$$\begin{aligned}
&\Pr((1,1)|Z) = \Pr(\epsilon_1 > -\alpha_1' Z_1 - \delta_1; \epsilon_2 > -\alpha_2' Z_2 - \delta_2) \\
&\Pr((0,0)|Z) = \Pr(\epsilon_1 \leq -\alpha_1' Z_1 - \delta_1; \epsilon_2 \leq -\alpha_2' Z_2 - \delta_2) \\
&\Pr((1,0)|Z) = \Pr((\epsilon_1,\epsilon_2) \in R_1(Z,\theta)) + \int \Pr(((1,0)|\epsilon_1,\epsilon_2,Z)\mathbf{1}\{(\epsilon_1,\epsilon_2) \in R_2(\theta,Z)\} dF_{\epsilon_1,\epsilon_2} \\
&\Pr((0,1)|Z) = 1 - \Pr((1,1)|Z) - \Pr((0,0)|Z) - \Pr((1,0)|Z)
\end{aligned}$$

where

$$\begin{aligned}
R_1(\theta,Z) = \{(\epsilon_1,\epsilon_2) &: (\epsilon_1 \geq -\alpha_1' Z_1; \epsilon_2 \leq -\alpha_2' Z_2) \\
&\cup (\epsilon_1 \geq -\alpha_1' Z_1 - \delta_1; -\alpha_2' Z_2 \leq \epsilon_2 \leq -\alpha_2' Z_2 - \delta_2)\} \\
R_2(\theta,Z) = \{(\epsilon_1,\epsilon_2) &: (-\alpha_1' Z_1 \leq \epsilon_1 \leq -\alpha_1' Z_1 - \delta_1; \\
&- \alpha_2' Z_2 \leq \epsilon_2 \leq -\alpha_2' Z_2 - \delta_2))
\end{aligned}$$

$Z = (Z_1, Z_2)$, $\theta = (\alpha', \delta', \sigma)'$ is a finite dimensional parameter of interests. For the first two equilibria, we can determine the ϵ - region where they uniquely exist. For equilibrium (1,0), R_1 is the region where it uniquely exists and R_2 is the region where it may coexists with (0,1). Note that $\Pr((1,0)|\epsilon_1,\epsilon_2,Z)$ specifies the probability when ϵ falls into the multiple equilibria region, but we have no information on this selection probability. However, we can derive bounds for the choice probability:

$$\Pr((\epsilon_1,\epsilon_2) \in R_1) \leq \Pr((1,0)) \leq \Pr((\epsilon_1,\epsilon_2) \in R_1) + \Pr((\epsilon_1,\epsilon_2) \in R_2)$$

These bounds can provide necessary conditions for partial identification.

3 Identification

We first make the following assumption:

Assumption 1: *We have a random sample of observations (y_m, Z_m), $m = 1, ..., n$. Let $n \to \infty$. Assume that ϵ is continuously distributed on $\mathbb{R}^2$ independently of $Z = (Z_1, Z_2)$ with a joint distribution $N(0, \sigma^2 I)$. The parameter of interest is $\theta = (\alpha', \delta', \sigma)'$.*

Then we can express the choice probability $y' \in Y := \{(0,0), (0,1), (1,0), (1,1)\}$ given Z:

$$
\begin{aligned}
\Pr(y'|Z) &= \int \Pr(y'|\epsilon, Z) dF_\epsilon \\
&= \int_{R_1(\theta,Z)} \Pr(y'|\epsilon, Z) dF_\epsilon + \int_{R_2(\theta,Z)} \Pr(y'|\epsilon, Z) dF_\epsilon \\
&= \underbrace{\int_{R_1(\theta,Z)} dF_\epsilon}_{\text{unique outcome region}} + \underbrace{\int_{R_2(\theta,Z)} \Pr(y'|\epsilon, Z) dF_\epsilon}_{\text{multiple outcome region}}
\end{aligned}
$$

Without the specifying the equilibrium selection function $\Pr(y'|\epsilon, Z)$, we exploit the fact that this probability is bounded between 0 and 1, and hence we have the following implication:

$$
\int_{R_1(\theta,Z)} dF_\epsilon \leq \Pr(y'|Z) \leq \int_{R_1(\theta,Z)} dF_\epsilon + \int_{R_2(\theta,Z)} dF_\epsilon
$$

for all $y' \in Y$. Write in a vector form:

$$
\mathbf{H_1}(\theta, \mathbf{Z}) := \begin{bmatrix} H_1^1(\theta, Z) \\ \vdots \\ H_1^4(\theta, Z) \end{bmatrix} \leq \begin{bmatrix} \Pr(y_1'|Z) \\ \vdots \\ \Pr(y_4'|Z) \end{bmatrix} \leq \begin{bmatrix} H_2^1(\theta, Z) \\ \vdots \\ H_2^4(\theta, Z) \end{bmatrix} =: \mathbf{H_2}(\theta, \mathbf{Z}) \qquad (*)
$$

where $\mathbf{H_1}(\theta, \mathbf{Z})$ is the lower bound of choice probabilities and $\mathbf{H_2}(\theta, \mathbf{Z})$ is the upper bound.

Notice that the middle column $\Pr(y'|Z)$ can be estimated from data, since Z has been discretized as follows: For each observation, $Z_m = (s_m, x_m, 1)$. Its market size and market presence s_m and x_{im} are assigned as 1 if it is above median and 0 otherwise. Therefore, Z can only take 8 values based on (s_m, x_{1m}, x_{2m}). This discretization has two advantages: First, it guarantees that we have finitely many moment inequalities, which is typically required by estimation methods. Second, it makes sure that our data have finitely many support points $z_i \in S_Z = \{z_1, ...z_8\}$. We can use a simple frequency estimator to approximate $\Pr(y'|Z)$:

$$\mathrm{P}_n(y'|Z) = \frac{\sum_i \mathbf{1}\{y_i = y'\}\mathbf{1}\{Z_i = Z\}}{\sum_i \mathbf{1}\{Z_i = Z\}}$$

and it is easy to prove the consistency:

$$\sup_z |\mathrm{P}_n(y'|Z) - \Pr(y'|Z)| = o_p(1)$$

This is a conditional moment inequality model and the identified feature is the set of parameter values that obey these restrictions for $\mathbf{Z}$ almost everywhere and the represents the set of economic models that is consistent with the empirical evidence. Formally, we give the definition of the identified set, which is our interest:

Definition 1: *Let Θ_I be the identified set such that*

$$\Theta_I = \{\theta \in \Theta \textit{ s.t. inequalities } (*) \textit{ are satisfied at } \theta \textit{ for } \forall Z \textit{ a.s.}\}$$

Θ_I is usually not a singleton. Many econometric literature has made several attempts to establish set consistent estimator $\hat{\Theta}_n$, but the condition is often not easily verifiable or are too restrictive. For example, Tamer (2003) has provided sufficient conditions that guarantee point identification given the exclusive restriction when the independent variable $z \to \infty$. However, even though a large variation in Z can indeed help shrink the set Θ_I, such conditions are usually too stringent. Alternatively, Chernozhukov, Hong and Tamer (2007) uses an idea to expand $\hat{\Theta}_n$ with a tolerance ϵ_n:

$$\hat{\Theta}_n^{\epsilon_n}(\theta) = \{\theta^* \in \Theta : \hat{Q}_n \leq \min_{\theta \in \Theta} \hat{Q}_n(\theta) + \epsilon_n\}$$

When $\Theta \subseteq [-B, B]^2$, letting $n\epsilon_n \to \infty$ guarantees consistency $d_H(\hat{\Theta}_n^{\epsilon_n}(\theta), \Theta_I) \to_p 0$, where d_H is the Hausdorff distance. However, the theory does not provide much guidance on the choice of ϵ_n. In this paper, we instead focus on the confidence set $CS_n(1-\alpha)$:

$$\Pr(\Theta_I \subseteq CS_n(1-\alpha)) \approx 1 - \alpha$$

From the confidence set, we can also obtain confidence intervals for each θ_k such that $\theta =$

$\prod_k \theta_k \in \Theta_I$.

4 Estimation

4.1 Conditional Chi-Squared Test

To construct the confidence set CS_n, I incorporate the moment inequality test method proposed by Cox and Shi (2023). This is a conditional chi-squared test (CC test) because it uses a critical value that is a quantile of the chi-squared distribution, where the degree of freedom depends on the active inequalities. Because the critical value is not obtained by simulation and no tuning parameter is required, this method is method is computationally fast. I now describe the procedure of the test under the settings of this binary airline entry game.

I. Construct the Moment Inequality Model

From above discussion, we can construct a moment inequality model as:

$$A\mathbb{E}[m(\mathbf{y}, Z)] \leq b(\theta)$$

where $\mathbb{E}[m(\mathbf{y}, Z)] = (\Pr(\mathbf{y}_1'|z_1)', \Pr(\mathbf{y}_2'|z_2)', ..., \Pr(\mathbf{y}_8'|z_8)')'$, and $\mathbf{y}_i \subseteq Y$ is a set contains possible equilibria which is specific to the support point z_i. Because the test statistics T_n involves computing the inverse of the covariance matrix of $\mathbb{E}[m]$, it should be linear independent otherwise the covariance matrix Σ would be singular.

Theoretically, one should set $\mathbf{y}_i = \{(0,0), (0,1), (1,0)\}$ or a similar "leave-one-out" subset to avoid linear dependency since the sum of probabilities of all equilibria $\sum_{j=1}^{4} \Pr(y_j|Z) = 1$, and the dimension of $\mathbb{E}[m(\mathbf{y}, Z)]$ should be $(4-1) \times 8 = 24$. However in practice, we need to adjust it according to z_i, since the four equilibria Y are not all observed for each z_i due to the limited sample size. The zero-probability issue will cause a problem because the corresponding variance will be zero thus the covariance matrix Σ is no longer invertible. Instead, I define $\mathbf{y}_i$ as a leave-one-out subset of Y_i, which is specific to z_i and contains all observable equilibria. The dimension of $\mathbb{E}[m(\mathbf{y}, Z)]$ being actually used is 13.

To specify the matrix A and $b(\theta)$, we first consider the case for one z_i and the corresponding

$n(Y_i) = k$. From the inequality $(*)$, we can write

$$\begin{bmatrix} \Pr(y_1'|z_i) \\ \vdots \\ \Pr(y_{k-1}'|z_i) \\ 1 - \sum_{j=1}^{k-1} \Pr(y_j'|z_i) \end{bmatrix} \leq \begin{bmatrix} H_2^1(\theta, z_i) \\ \vdots \\ H_2^{k-1}(\theta, z_i) \\ H_2^k(\theta, z_i) \end{bmatrix}$$

for the upper bound $H_2(\theta, z_i)$ and

$$\begin{bmatrix} -\Pr(y_1'|z_i) \\ \vdots \\ -\Pr(y_{k-1}'|z_i) \\ \sum_{j=1}^{k-1} \Pr(y_j'|z_i) - 1 \end{bmatrix} \leq \begin{bmatrix} -H_1^1(\theta, z_i) \\ \vdots \\ -H_1^{k-1}(\theta, z_i) \\ -H_1^k(\theta, z_i) \end{bmatrix}$$

for the lower bound $H_1(\theta, z_i)$. Combine this two, we get

$$\underbrace{\begin{bmatrix} I_{k-1} \\ -\mathbf{1} \\ -I_{k-1} \\ \mathbf{1} \end{bmatrix}}_{A_i} \underbrace{\begin{bmatrix} \Pr(y_1'|z_i) \\ \vdots \\ \Pr(y_{k-1}'|z_i) \end{bmatrix}}_{\mathbb{E}[m(\mathbf{y}_i, z_i)]} \leq \underbrace{\begin{bmatrix} H_2^1(\theta, z_i) \\ \vdots \\ H_2^{k-1}(\theta, z_i) \\ H_2^k(\theta, z_i) - 1 \\ -H_1^1(\theta, z_i) \\ \vdots \\ -H_1^{k-1}(\theta, z_i) \\ 1 - H_1^k(\theta, z_i) \end{bmatrix}}_{b_i(\theta)}$$

which has the form $A_i \mathbb{E}[m(\mathbf{y}_i, z_i)] \leq b_i(\theta)$. Since this inequality has to hold for all $z_i \in S_Z$. Therefore, we define A as the matrix which is formed by stacking A_i $(i = 1, ..., 8)$ on the diagonal, and $b(\theta) := [b_1(\theta), ..., b_8(\theta)]'$. Along with the previously defined $\mathbb{E}[m(\mathbf{y}, Z)]$, we now have formed the desired moment inequality model.

Notice that, when all equilibria are observed in the data and no zero-probability issue presents, the number of rows of A should be $(4 + 4) \times 8 = 64$ and the number of columns should be

$(3-1) \times 8 = 24$. However, in practice the dimension of A is 42×13 because some equilibria are never observed for some z_i thus I need to exclude them from the sample space.

II. Simulation for $b(\theta)$

The model structure makes it impossible to derive $\mathbf{H_1}$ and $\mathbf{H_2}$ analytically. Suggested by McFadden (1989) and Pakes and Pollard (1989), We use simulation to construct $b(\theta)$. However, due to the zero-probability issue, some modifications are necessary. Given one θ value, $b(\theta)$ is formed following steps below:

1. Define the profit function for both players:

$$\pi_j(\theta, y, z_i) = \alpha_{j1}s_i + \alpha_{j2}x_i + \alpha_{j3} + \delta_j y_{3-j} + \epsilon_j, \quad j = 1, 2$$

2. For each simulation $s = 1, 2, ...$, draw $\epsilon = (\epsilon_1, \epsilon_2)$ from $N(0, \sigma^2 I)$

3. For each equilibrium $y_i \in Y_i$, compute the realization of player's profits (π_1, π_2). If $y_i = (\mathbf{1}\{\pi_1 > 0\}, \mathbf{1}\{\pi_2 > 0\})$, store y_i into the equilibrium set Y_i^s

* *Caveat: If no equilibrium y_i is stored for some simulation, we consider this simulation is invalid and redraw the ϵ*

4. For each y_i, if $y_i \in Y_i^s$, $\hat{H}_2(y_i|z_i) = \hat{H}_2(y_i|z_i) + 1$. Furthermore, if $y_i = Y_i^s$ is a singleton, we also do $\hat{H}_1(y_i|z_i) = \hat{H}_1(y_i|z_i) + 1$

5. Repeat step 2-4 until $s = S$

6. Divide $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{H}}_2$ by simulation times: $\hat{\mathbf{H}}_1 = \frac{1}{S}\hat{\mathbf{H}}_1$, $\hat{\mathbf{H}}_2 = \frac{1}{S}\hat{\mathbf{H}}_2$

7. Construct $b(\theta)$ use the expression stated in previous section

The reason for excluding invalid simulation is because Y_i is not necessary equal to Y that includes all 4 possible equilibria. Since some equilibria may not observed for z_i, we need to shrink the sample space accordingly.

III. Procedure of the CC Test

We conduct inference for the true parameter Θ_I by test inversion: Given a significance level $\alpha \in (0, 1)$, we construct a test $\phi_n(\theta, \alpha)$ for $H_0 : \theta \in \Theta_I$, where $\phi_n(\theta, \alpha) = 1$ indicates rejection and $\phi_n(\theta, \alpha) = 0$ indicates failure to reject. First we need to compute the test statistic $T_n(\theta)$, which is a (quasi-) likelihood ratio:

$$T_n(\theta) = \min_{A\mu \leq b(\theta)} (m(\mathbf{y}, Z) - \mu)' \hat{\Sigma}^{-1} (m(\mathbf{y}, Z) - \mu)$$

where $\hat{\Sigma}$ is the estimated covariance matrix of $\hat{m}(Z, \theta)$ and it can be obtained by bootstrap.

The test uses a data-dependent critical value that is based on the rank of the rows of A corresponding to the inequalities that are active in finite sample. To define it rigorously, let $\hat{\mu} = \text{argmin}_{A\mu \leq b(\theta)} (m(Z, \theta) - \mu)' \hat{\Sigma}^{-1} (m(Z, \theta) - \mu)$. This is a restricted estimator for the moments which can be calculated using a quadratic programming algorithm. Let a_j' denote the j-th row of A and $b_j(\theta)$ denote the j-th element of $b(\theta)$. We define

$$\hat{J} = \{j \in \{1, 2, ..., d_A\} : a_j' \hat{\mu} = b_j\}$$

which is the set of indices for the active inequalities.

For a set $J \subseteq \{1, 2, ...d_A\}$, let A_J be the submatrix of A formed by the rows of A corresponding to the element of J. Let $\hat{r} = \text{rank}(A_{\hat{J}})$. However, due to the way A is constructed, we notice that $\hat{r} = \text{rank}(A_{\hat{J}}) = n(\hat{J})$, where $n(\hat{J})$ is the number of elements in $\hat{J}$. This simplification will speed up the computation. We should also notice that $\hat{\mu}$, $\hat{J}$ and $\hat{r}$ need to be calculated for every θ value for the test inversion.

The critical value $\text{cv}(\theta)$ of the CC test is the $100(1-\alpha)$ quantile of $\chi^2_{\hat{r}}$, the chi-squared distribution with $\hat{r}$ degrees of freedom, denoted by $\chi^2_{\hat{r},1-\alpha}$. Then, the CC test is defined as:

$$\phi_n^{CC}(\theta, \alpha) = \mathbf{1}\{T_n(\theta) > \chi^2_{\hat{r},1-\alpha}\}$$

Then we can define the confidence set under the setup of the CC test:

Definition 2: The confidence set $CS_n(1-\alpha)$ given by the CC test is

$$CS_n(1-\alpha) = \{\theta \in \Theta : \phi_n^{CC}(\theta, \alpha) = 0\} = \{\theta \in \Theta : T_n(\theta) < \chi^2_{\hat{r}(\theta),1-\alpha}\}$$

4.2 Construct Confidence Intervals

We want to construct confidence intervals for all $\theta_k = \pi_k(\theta)$ for $\forall\theta \in CS_n(1-\alpha)$, where $\pi_k : \theta \to \theta_k$ is the coordinate map and $\theta = \prod_k \theta_k$. The simplest idea is to conduct grid search for all θ_k and pick up the upper/lower bounds. However, θ is 9-dimensional: 4 parameters (α_{i1},α_{i2},α_{i3},δ_i) for each player i corresponding to market size, market presence, constant and the opponent's entry, and σ for the error term. The curse of dimensionality will cause the pure grid search computationally infeasible. Instead, I incorporate the CC test and define the following objective functions:

$$Q_n^{ub}(\theta_k) = \max\{\theta_k : \inf_{\theta_{-k}\in\Theta_{-k}} [T_n(\theta_{k,-k}) - \text{cv}(\theta)] < 0\}$$

$$Q_n^{lb}(\theta_k) = \min\{\theta_k : \inf_{\theta_{-k}\in\Theta_{-k}} [T_n(\theta_{k,-k}) - \text{cv}(\theta)] < 0\}$$

The main idea is, we want to find an confidence interval for θ_k such that for every element within, there exist the rest of the components θ_{-k} that guarantees $\theta_{k,-k}$ cannot be rejected. Then we define Q_n^{ub} and Q_n^{lb} as the upper bound and lower bound of this confidence interval.

In practice, I combine the minimization and grid search following the steps below:

1. Staring from the first parameter θ_1, conduct grid search on $[\underline{\theta}_1, \bar{\theta}_1]$ with step size $(\bar{\theta}_1 - \underline{\theta}_1)/N$

2. For each value in the grid $\theta_1[i]$, minimize the objective:

$$\min_{\theta_{-1}} T_n((\theta_1[i], \theta_{-i})) - \text{cv}((\theta_1[i], \theta_{-i}))$$

3. If the objective function value is negative, accept $\theta_1[i]$ and store the argmin $\theta = (\theta_1[i], \theta_{-1}^*)$ to the confidence region $CS_n(1-\alpha)$, otherwise reject

4. Repeat the step 1-3 of grid search for all θ_k

5. From the confidence region already acquired, construct the confidence interval for each θ_k by

$$\theta_k^{ub} = \max\{\theta_k : \exists\theta_{-k} \text{ s.t. } (\theta_k, \theta_{-k}) \in CS_n(1-\alpha)\}$$

$$\theta_k^{lb} = \min\{\theta_k : \exists\theta_{-k} \text{ s.t. } (\theta_k, \theta_{-k}) \in CS_n(1-\alpha)\}$$

This method is free from the dimensionality curse since for N points in each parameter grid, only $9N$ points need to be evaluated.

4.3 Non-Convexity of The Objective Function

One problem of our previous method is the minimization of objective $T_n(\theta) - \text{cv}(\theta)$ is non-smooth and non-convex. The solver could easily get stuck in a local minimum, which causes an over-rejection issue. As a consequence, our estimator for confidence intervals will be conservative.

One commonly used solution is to randomly simulate M starting parameters to repeat the minimization M times and then select the minimum among them (in practice, I set $M = 30$). If M is sufficiently large, it will cover all parameter space and local minima are avoided. Moreover, if $M \to +\infty$, this method is equivalent to finest grid search. However, I notice that increasing M not only significantly slow down and computation but also the efficiency is unsatisfied. Therefore, I develop methods that effectively alleviate the over-rejection problem. These methods can not only applied in this specific problem but may also helpful in other areas which involves optimize over a non-convex objective function. I will elaborate in the following sections:

4.3.1 Strategic Selection

To put it briefly, this method involves strategically select a starting parameter which is "closest" to our testing value from the confidence region. The intuition is, when we test $\theta_k[i]$ given that we already know a close value $\theta_k[j]$ is not rejected, $(\theta_k[j], \theta^*_{-k})$ would serve as a good starting value for $\theta_k[i]$. In practice, I follow the steps below:

1. I search manually for one parameter θ^0 that is not rejected by CC test and store it in the confidence region $CS_n(1-\alpha)$
2. When doing grid search for θ_1 and test over $[\underline{\theta}_1, \bar{\theta}_1]$, select the closest $\tilde{\theta} \in CS_n$ such that $|\theta_1[i] - \tilde{\theta}_1| = \min\{|\theta_1[i] - \theta_1| : \theta \in CS_n\}$. (Since CS_n is only a singleton now, simple use θ^0 as the starting point)
3. If $\theta_1[i]$ is accepted, store $(\theta_1[i], \theta^*_{-1})$ into CS_n
4. Repeat this process for all θ_k

As the test goes on, CS_n will be expanded and be able to provide more helpful starting points. One can make modifications on this selection process to improve efficiency, such as starting the test not straightly from $[\underline{\theta}_1, \bar{\theta}_1]$ but from the value of $\theta_1^0 + \epsilon$, since a closer starting parameter will reduce the probability of the type I error especially when CS_n is still small. Alternatively, one can select the closest n parameters from the confidence region instead of just one.

I report the estimated confidence intervals using the default method (one starting parameter θ^0 and M random guesses) in column 1 of Table 5, and the second method (one starting parameter θ^0, one strategically selected $\tilde{\theta}$ and M random guesses) in column 2. The results show this method significantly reduce the conservatism of the estimation.

4.3.2 Seed Control

Another way to improve the efficiency of starting parameters is to optimize our M guesses. Intuitively, one should not use seeds for these simulations, since randomness is the key to cover a broader parameter space. However, I found that during each round of evaluation, most of the guesses are not helpful since they deliver too large objective values. The idea of this method is to record the effectiveness of M simulations, and only keep the good ones. I achieve this goal by using seeds to control the simulation:

1. Use one (hyper)seed to generate M random positive integers which will be used as seeds for simulation
2. Generate M starting parameters using seeds
3. For each parameter value $\theta_k[i]$, repeat the optimization for each starting parameter and record the objective value
4. If $\theta_k[i]$ is rejected for all starting parameters, update seeds for simulation by the following rule:
    (a) Keep the $M/2$ seeds that correspond to the lowest $M/2$ objective values
    (b) Random generate the other $M/2$ integers as seeds (note that for randomness, we should not use seed to control this part)
    (c) Combine the two halves and use it to generate M new starting parameters for the next parameter value $\theta_k[i+1]$ to be tested

5. If If $\theta_k[i]$ is accepted, skip this updating process

The above procedure will guarantee that whenever a rejection happens, "fresh water" will be introduced to replace bad guesses. One can also modify this idea by choosing different updating proportion $M/2$; or when the rejection happens, update seeds and redo the evaluation K times. I report the results incorporating both strategic selection and seed control in column 3 of Table 5. Results show that this method also leads to a significant improvement.

One caveat is that, using seed control does not guarantee to improve the performance every time, even though it usually does. When introducing $M/2$ new seeds, randomness is still retained so it is possible that we still get bad luck. Practically, one can repeat the estimation several times and take the union of confidence intervals to achieve the best results.

5 Data

5.1 Data Construction

My data is collected from the first quarter of the 2019 Airline Origin and Destination Survey (DB1B) and the T-100 airline database from the U.S. Department of Transportation. The data construction is similar to Ciliberto and Tamer (2009). The market is defined as a trip between two airports regardless of intermediate transfer points and of the direction of the flight. My data contains 1388 observed markets in total including both hub markets and non-hub markets. There are 5 airlines operating in this period: American Airlines (AA), Delta Airlines (DL), United Airlines (UA), Southwest Airline (WN) and other low cost carriers (LCC). Among them, AA, DL and UA are defined as major airlines and I construct hub markets based on whether a market has an endpoint city that serves as a hub for any major airline. According to Wikipedia, hubs for 3 major airlines are reported in Table 1. Based on the hub information, my data has 886 recorded hub market.

Market size variable is defined as the population at the endpoints of the market. I merge the data in Ciliberto and Tamer (2009) to mine and obtain 645 matched observations. Moreover, I compute the market presence variable for every airline as follows: For each market, I compute the number of markets that the airline serves from one airport and divided by the total number of markets served from that airport by any airline, then averaging between two endpoint airports.

After creating market size and market presence, I discretize these variables by comparing to its median (0 if below the median and 1 otherwise). Hence, the actual data used in my analysis are all binary. One can use a finer partition to obtain a sharper estimation, as what Ciliberto and Tamer (2009) does is discretize according to 0, 25, 50, 75 quantiles. However, this will greatly increase the number of moment inequalities using in my method since the space S_Z is expanded.

5.2 Descriptive Statistics

To explore the data structure, I first report the conditional entry probability for these two players in Table 2. Overall, non-hub player enters a larger proportion of the markets (85.89%) comparing to hub player (56.28%), since by definition non-hub player includes more airlines comparing to hub player. The results show both have larger probabilities of entry when market size and market presence of itself are above median. Their profits clearly benefit from these characteristics. Moreover, they both show clearly negative competitive effects since the conditional entry probabilities are lower when the opponent is present. Notably, hub player's entry probability declines sharper (from 98.90% to 49.28%) than non-hub player (from 99.65% to 75.21%).

Then, I conduct probit and linear probability regressions for their entries as reported in Table 3 and Table 4, respectively. The results make it possible to compare two players' reactions clearly. Both regressions show the similar pattern: For market size, their reactions are of the similar scale, but hub player benefits slightly more. For market presence, hub player shows an obvious advantage (1.664 comparing to 0.385 in probit regression, 0.574 comparing to 0.087 in LP regression), which make sense because a hub presence is strongly correlated to an airline's network. The advantage of the hub will encourage hub player to enter much more frequently than non-hub player. For the competitive effect, the result is different from the previous conditional probability observation, since non-hub player demonstrates a larger sensitivity to the opponent's entry. For the probit regression, the coefficient for competition effect of hub player is -1.941 and -2.069 for non-hub player; while in LP regression, the coefficient is -0.228 for hub player and -0.247 for non-hub player.

6 Results

I report the estimated confidence intervals in Table 5. The results are in 2 decimal points for computational simplicity. Results in the first column are estimated using $M+1$ starting parameters ($M = 30$ in practice), where M are randomly simulated from a predetermined bounds. I construct this predetermined bounds using an iterative process: First, I construct wide bounds that are reasonable for all parameters, then use it for estimation. The estimation results will narrow the bounds so that I can eliminate some regions that are never accepted. Then, I use this updated bounds to do the estimation again. Additionally, I manually find one parameter θ^0 that is not rejected and include it in my starting parameter bucket since it can help to find other nearby points in the confidence region. The results show that the confidence intervals are approximately centered around θ^0 and are much narrower since over-rejection causes estimation to be conservative.

The second column contains results using strategic selection. Among $M+1+1$ starting parameters, M are still randomly simulated and θ^0 is still included (to guarantee a better performance than the previous case), but I select an additional starting parameter strategically from the confidence region as discussed in Section 4.3.1. The estimated confidence intervals are wider than previous case and show similar patterns.

The third column contains results using both strategic selection and seed control. Now M of them are controlled by seeds to record their performance when testing each point in the grid. Comparing to the second column, most of the confidence intervals are wider, but there are some minor exceptions: As I have discussed in Section 4.3.2, seed control has to retain randomness so it does not guarantee to have a better performance for every parameter and every bound, though it performs better overall.

From the estimated intervals, one can get a clearer sense of the game's payoffs. Both players react similarly to market size (but hub player is slightly better), and hub player benefit significantly more from market presence ($[1.11, 2.85]$ comparing to $[-0.12, 0.90]$). Moreover, hub player demonstrates more robustness to competition ($[-2.50, -1.61]$ versus $[-3.00, -1.12]$), though all the regions are clearly negative. Additionally, hub player's payoff relies more heavily on market characteristics, since when both market size and market presence are zeros and no opponent enters, the constant represents a "default" payoff, which is much smaller for hub player ($[-0.71, 1.32]$) than non-hub player ($[1.34, 3.23]$).

7 Conclusion

In this paper, I adopt the classic framework from Ciliberto and Tamer (2009) and construct a binary entry game to study airlines' entry decision in hub markets. This framework allows for a general form of heterogeneity across players without making equilibrium selection assumptions and focus only on partial identification. In the simple game I constructed, hub player and non-hub player's payoffs depend on market size, their own market presence, and the opponent's entry. I assume they make entry decision simultaneously, thus when the random shock ϵ falls into the middle region, multiple equilibria may appear.

To deal with multiple equilibria, I construct upper and lower bounds for these equilibria and form a moment inequality model which can be tested by the method proposed by Cox and Shi (2023), which possesses advantages of requiring no tuning parameter, no simulation and computationally fast. I furthermore incorporate this test to develop a novel way to estimate confidence intervals different from Chernozhukov, Hong and Tamer (2007). My method combines grid search and optimization so that it is free from the curse of dimensionality and is open to flexible modifications. Moreover, since the objective function is non-smooth and non-convex, which is also encountered by many other methods, the estimated intervals are often conservative. I invent two strategies to alleviate this problem by exploit the structure of the confidence region: First, when testing each point in the grid, I strategically select the "closest" parameter from the confidence region aiming to enhance efficiency. As the test goes on, I store the accepted parameters into the confidence region so that it will expand over time and be able to provide better starting parameters. Second, I use seeds to record performance of simulated starting parameters and control the simulation by introducing new random guesses to replace bad ones whenever rejection happens. My results show both strategies perform well comparing to random simulation.

My estimated confidence intervals well demonstrate the two player's payoff structures in this game. Their payoffs react similarly to market size, though hub player benefits slightly more. For market presence, hub player shows a significant advantage comparing to non-hub player and benefit much more from the network effect. Moreover, hub player are less prone to the opponent's entry since its payoff are less negatively affected by competition effect. Overall, hub player relies more heavily on market characteristics than non-hub player, since the "default" payoff are much lower.

References

[1] S. T. Berry. Estimation of a model of entry in the airline industry. *Econometrica: Journal of the Econometric Society*, pages 889–917, 1992.

[2] P. A. Bjorn and Q. H. Vuong. Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation. 1984.

[3] T. F. Bresnahan and P. C. Reiss. Entry and competition in concentrated markets. *Journal of Political Economy*, 99(5):977–1009, 1991.

[4] V. Chernozhukov, H. Hong, and E. Tamer. Estimation and confidence regions for parameter sets in econometric models 1. *Econometrica*, 75(5):1243–1284, 2007.

[5] F. Ciliberto and E. Tamer. Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828, 2009.

[6] G. Cox and X. Shi. Simple adaptive size-exact testing for full-vector and subvector inference in moment inequality models. *The Review of Economic Studies*, 90(1):201–228, 2023.

[7] P. Jia. What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, 76(6):1263–1316, 2008.

[8] B. Kline and E. Tamer. Bayesian inference in a class of partially identified models. *Quantitative Economics*, 7(2):329–366, 2016.

[9] D. McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, pages 995–1026, 1989.

[10] A. Pakes and D. Pollard. Simulation and the asymptotics of optimization estimators. *Econometrica: Journal of the Econometric Society*, pages 1027–1057, 1989.

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

Figures and Tables

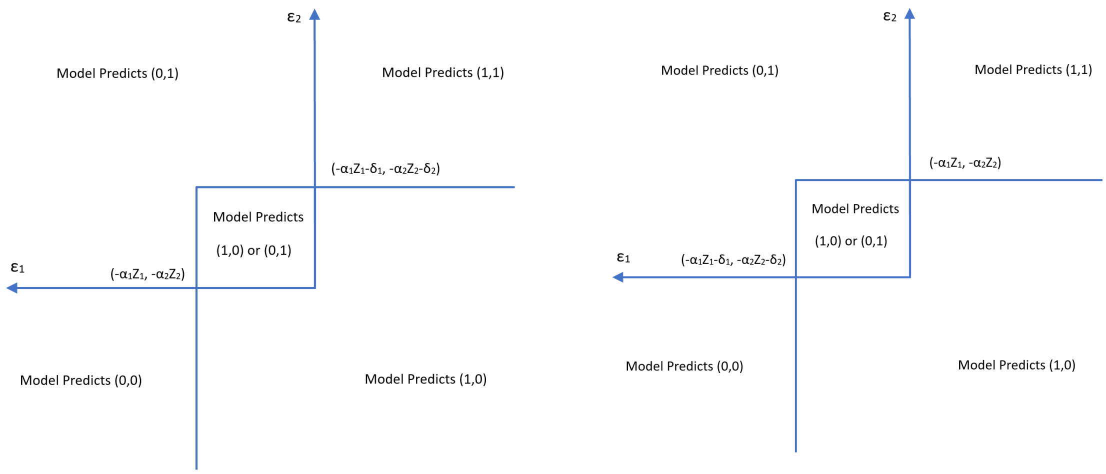

Figure 1: Regions for multiple equilibria: LHP, $\delta_1, \delta_2 < 0$; RHP, $\delta_1, \delta_2 > 0$

Airline	**Hubs**
Delta Airlines (DL)	Atlanta (ATL), Boston (BOS), Detroit (DTW), Los Angeles (LAX), Minneapolis (MSP), New York (JFK), New York (LGA), Salt Lake City (SLC), Seattle (SEA)
United Airlines (UA)	Chicago (ORD), Denver (DEN), Guam (GUM), Houston (IAH), Los Angeles (LAX), Newark (EWR), San Francisco (SFO), Washington-Dulles (IAD)
American Airlines (AA)	Charlotte (CLT), Chicago (ORD), Dallas/Fort Worth (DFW), Los Angeles (LAX), Miami (MIA), New York (JFK), New York (LGA), Philadelphia (PHL), Phoenix Sky Harbor (PHX), Washington National (DCA)

Table 1: Hubs of U.S. Major Airlines

Condition/Player	Hub Player	Non-Hub Player
all markets	56.28%	85.89%
market size = 1	60.56%	88.51%
market size = 0	52.01%	83.28%
market presence of hub player = 1	83.89%	75.28%
market presence of hub player = 0	21.40%	99.30%
market presence of non-hub player = 1	54.97%	88.58%
market presence of non-hub player = 0	59.88%	78.49%
hub player entry = 1	100.00%	75.21%
hub player entry = 0	0.00%	99.65%
non-hub player entry = 1	49.28%	100.00%
non-hub player entry = 0	98.90%	0.00%

Table 2: Conditional Entry Probabilities For Two Players

Regressor	Hub Player	Non-Hub Player
market size	0.490***	0.369**
	(0.123)	(0.144)
market presence of hub player	1.664***	-
	(0.122)	
market presence of non-hub player	-	0.385**
		(0.153)
non-hub player entry	-1.941***	-
	(0.510)	
hub player entry	-	-2.069***
		(0.347)
constant	0.847*	2.300***
	(0.511)	(0.352)

Note: $^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$; z statistics in parentheses

Table 3: Probit Regression Results

Regressor	Hub Player	Non-Hub Player
market size	0.118***	0.071***
	(0.030)	(0.026)
market presence of hub player	0.574***	-
	(0.032)	
market presence of non-hub player	-	0.087***
		(0.029)
non-hub player entry	-0.228***	-
	(0.046)	
hub player entry	-	-0.247***
		(0.026)
constant	0.379***	0.899***
	(0.052)	(0.031)

Note: $^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$; t statistics in parentheses

Table 4: Linear Probability Regression Results

	Single Starting Parameter		Strategic Selection		Strategic Selection and Seed Control	
	Hub Player	**Non-Hub Player**	**Hub Player**	**Non-Hub Player**	**Hub Player**	**Non-Hub Player**
market size	[0.05, 0.62]	[-0.40, 0.68]	[-0.04, 0.87]	[-0.4, 0.89]	[-0.04, 1.01]	[-0.4, 0.97]
market presence	[1.11, 2.03]	[0.12, 0.90]	[1.11, 2.49]	[0.00, 1.17]	[1.11, 2.85]	[-0.12, 0.90]
competitive effect	[-2.07, -1.82]	[-2.85, -2.13]	[-2.3, -1.73]	[-2.86, -1.12]	[-2.50, -1.61]	[-3.00, -1.25]
constant	[0.82, 1.07]	[1.8, 2.52]	[0.30, 1.34]	[1.32, 3.00]	[-0.59, 1.49]	[1.34, 3.23]
error term ϵ	[0.89, 1.25]		[0.71, 1.32]		[0.80, 1.50]	

Table 5: Estimated Confidence Intervals