Identification and Estimation of Incomplete Information
Games with Multiple Equilibria*

Ruli Xiao
Johns Hopkins University
November 20, 2013

Job Market Paper

Abstract

The presence of multiple equilibria in games is a big challenge for identification and estimation. Without information of the equilibrium selection, it is impossible to perform counterfactual analysis. Allowing for possibly multiple equilibria, this paper provides nonparametric identification of finite games with incomplete information. Upon observing players’ actions from cross-sectional games, the identification is achieved in two steps. First I identify the equilibrium-specific components, such as the number of equilibria, the equilibrium selection mechanism, and all strategies associated with each equilibrium. In particular, the econometric structure resembles that in measurement error models if I index the underlying equilibria and treat the index as a latent variable, and therefore, these equilibrium-specific components can be identified using results from measurement error literature. Next I identify the payoff functions with exclusion restrictions. This paper proves identification for both static and dynamic settings. Specifically, identification in static games requires at least three players with the standard assumption of independent private shocks, while identification in dynamic games requires four periods of data if only Markov Perfect Equilibria are considered. I apply this methodology to study the strategic interaction among radio stations when choosing different time slots to air commercials. The empirical results show the existence of two equilibria in smaller markets, and these markets exhibit the same equilibrium over time.

JEL Classification: C14

Keywords: Multiple equilibria, Discrete games, Measurement error models, Nonparametric Identification, Semiparametric estimation

*I am deeply indebted to Yingyao Hu for generous support and guidance. I benefited greatly from the comments of Yuya Sasaki and Richard Spady. I thank Andrew Sweeting, Victor Aguirregabiria, Jorge Balat, Hulya Eraslan and seminar participants at Johns Hopkins for their helpful comments. The usual disclaimer applies. Comments are welcome.
1 Introduction

Unlike single-agent discrete choice models, games generally admit multiple equilibria. Ignoring multiplicity in games may result in mis-specification and make it impossible to infer policy effects without the information of the equilibrium selection. To avoid mis-specification while enabling counterfactual analysis, this paper provides a methodology to identify game primitives and equilibrium-specific components nonparametrically for finite games with incomplete information. The methodology of identification imposes no restrictions on the cardinality of the equilibrium set or the equilibrium selection rules.

This paper studies games of incomplete information with finite actions. In such games, players receive random shocks of payoff before deciding their actions. Those payoff shocks are assumed to be private information, while the distributions of those shocks are common knowledge. The cardinality of the equilibrium set for such games is unknown but discrete and finite, so I can index the equilibria. Treating the equilibrium index as a latent variable, this paper provides a methodology to identify all equilibrium-specific components using results from measurement error literature. To begin with, I identify the number of equilibria. Next I identify the equilibrium selection mechanism and all players strategies in each equilibrium. Then I identify payoff primitives following the standard approach with exclusion restrictions. The identification procedure is constructive so that an estimator follows naturally. Applying this methodology, I study the strategic interaction among radio stations when choosing to air commercials during two time slots.

My methodology contributes to the literature on identification and estimation in games with multiplicity. Firstly, the methodology connects the identification of games with that of measurement error models, suggesting a new direction for identification of other games with possibly multiple equilibria. Secondly, the methodology nonparametrically identifies and estimates the number of equilibria, which is important because multiple equilibria are useful to explain important aspects of economic data. Thirdly, the methodology nonparametrically identifies and estimates the equilibrium selection, which reduces concern about model mis-specification and enables conduct counterfactual analysis. Moreover, given no guide from theory on how equilibria being selected, the estimated equilibrium selection sheds light on our understanding in the field. Additionally, I can link equilibrium characteristics with the selected rules, as in Bajari, Hong, and Ryan (2010). Lastly, the methodology provides an easily imple-

---

1 Examples of such games include static entry games in Bresnahan and Reiss (1990) and dynamic music format repositioning in Sweeting (2011)

2 For example, local market variation in firms strategic behavior (Sweeting (2009) and Ellickson and Misra (2008))
mented and computationally convenient method to estimate game primitives. As is widely known, computing all equilibria using Homotopy method is costly (Bajari, Hong, Krainer, and Nekipelov (2010a)), while the estimation in this paper does not need to solve for even a single equilibrium.

The methodology of identification applies to both static and dynamic games. In static games with cross-sectional data, identification relies on the standard assumption that private shocks are independent across actions and players. This assumption implies that actions of players are independent across different games when the same equilibrium is selected in those games. Given that players’ actions are independent conditional on the equilibrium index, I can use players’ actions as measurements for the equilibrium index. Recovering the cardinality of the index requires sufficient variation of each measurement. Moreover, identifying distributions of players’ actions for each equilibrium requires that games have at least three players. In static games with panel data, I relax the conventional assumption that the same equilibrium selected over time by allowing the equilibrium index to follow a first-order Markov process. Identification relies on the correlation between actions of all players across time.

As for dynamic games, I achieve the identification for Markov Perfect Equilibria (MPE) concept. The Markovian assumption implies that a single equilibrium is played in a market-level time series (Pesendorfer and Schmidt-Dengler (2008)). With MPE, the state, defined as all observables in each period including the game characteristic and actions by all players, follows a first-order Markov process for an individual game. In the presence of multiple equilibria, the state law of motion for the Markov process varies for games with different equilibria. Treating the equilibrium index as a latent variable, the Markovian structure with multiple equilibria resembles that with unobserved state variables in Hu and Shum (2012). The correlation between states over time reveals the aggregation of equilibrium-specific state law of motions. Thus, I prove that three periods of data is sufficient to identify the number of equilibria and that four periods of data are needed to identify the equilibrium-specific law of motion and the equilibrium selection rules.

Applying the proposed methodology to the data on radio stations studied in Sweeting (2006), I study the strategic interactions of stations when they choose to air commercials during two different time slots. The interaction is captured by a game with incomplete information, which fits the setup of identification. Treating market each period as an independent market, the estimation results show that smaller markets indeed admit two equilibria, which supports the conjecture of the existence of multiple equilibria, and this result is consistent with the findings in ?. By investigating the panel structure of the data, I find that smaller markets exhibit the same equilibrium over time. This finding supports the
conventional assumption that players select the same equilibrium overtime.

The remainder of the paper is organized as follows. Section 2 surveys related literature. Section 3 outlines the static game framework, and provides the nonparametric identification of the game. Section 4 describes the estimation procedure for static games. Section 5 provides the nonparametric identification results for dynamic games, followed by a Monte Carlo illustration to provide evidence for the methodology in section 6. Then I provide an empirical application to radio station commercial airing in section 7. Section 8 concludes. The Appendix contains the proofs, the figures and the tables.

2 Literature Review

This paper is related to recent literature on the econometric analysis of games in which multiple solutions are possible and identification with unobserved heterogeneity. In this section I summarize some of the recent findings.

Literature utilize different techniques to deal with identification and estimation of games in which multiplicity is possible. See De Paula (2012) for a survey of the recent literature on the econometric analysis of games with multiplicity. The information structure is an important guide for the econometric analysis, and different methods are developed for complete and incomplete information games. In incomplete information games, researcher usually assume that a unique equilibrium is selected in the data, which guarantees consistent estimation of the game primitives. See, e.g., Seim (2006) and Aradillas-Lopez (2010). One can be agnostic about the equilibrium selection rule. This is because all equilibria satisfy the same equilibrium conditions. Moreover, it is computationally challenging to compute all the equilibria through Homotopy method. Even though the degenerated equilibrium selection assumption guarantees consistent estimation of game primitives, it is nearly impossible to simulate the model and provide counterfactual inference. Moreover, the unique equilibrium assumption lacks both theoretical and empirical supports.

On the other hand, there are various approaches to inference in games of complete information games because it is easier to compute all the equilibria for any given model configurations. With all equilibria being computed, one approach is to focus on certain quantities that are invariant across different equilibria when more than one equilibrium is possible. See Berry (1992), Bresnahan and Reiss (1990) and Bresnahan and Reiss (1991). The key insight of this approach is that certain outcomes can only occur as a unique equilibrium, which limits the scope of its application. However, in some
settings, any outcomes bundle together in different equilibria. Tamer (2003) nevertheless identifies game primitives using exclusion restrictions and large support conditions on the observable covariates. The identification results rely on the extreme values of covariates and reduces decisions in games to a single-agent decision problem. However, this identification-at-infinity strategy leads to a slower asymptotic convergence rate (Khan and Tamer (2010)). Instead of relying on identification-at-infinity, another strand of literature use bound estimation instead of point estimation, relying on inequalities created by multiple equilibria\(^3\). Bajari, Hong, and Ryan (2010) incorporate a parameterized equilibrium selection function into the problem\(^4\), and identification is demonstrated with large support of the exclusion restrictions. Unlike Bajari, Chernozhukov, Hong, and Nekipelov (2009), this paper considers incomplete information games, and the equilibrium selection is nonparametrically recovered.

Another strand of literature focuses on testing or taking advantage of the presence of multiple equilibria. Sweeting (2009) points out that multiplicity helps for the identification of payoff primitives by providing additional information. Instead of attempting to identify the payoff primitives, De Paula and Tang (2012) use the fact that players’ equilibrium choice probabilities move in the same direction. As a result, the presence of multiplicity helps for identification of the sign of the interaction term. Echenique and Komunjer (2009) test complementarities between continuous explanatory and dependent variables in models with multiple equilibria.

This paper is related to current literature on identification with unobserved heterogeneity\(^5\). In particular, the identification intuition in static games follows identification results of misclassification errors in Hu (2008). In dynamic frameworks, Kasahara and Shimotsu (2009) consider the identification of dynamic discrete choice models with agents of a finite number of types, and demonstrate that the Markov law of motion is identified using six periods of data. In contrast, four periods of data is sufficient for identification in this paper. The identification is similar to that of dynamic models in Hu and Shum (2012) with continuous unobserved state variables, which are allowed to vary over time. The identification is also established using four periods of data. However, I study dynamic finite games with considering Markov Perfect Equilibria, in which the number of equilibria and actions are both discrete. Furthermore, the number of equilibria is endogenous in games and needed to be identified first.

\(^3\)Bounds estimation has also been used by Ciliberto and Tamer (2009), Pakes, Porter, Ho, and Ishii (2006), and Andrews, Berry, and Jia (2004). Berry and Tamer (2006) and Berry and Reiss (2007) survey the econometric analysis of discrete games.

\(^4\)See also Bjorn and Vuong (1984) and Ackerberg and Gowrisankaran (2006).

The paper is also related to Aguirregabiria and Mira (2013) (hereafter AM), which considers identification of incomplete information games with both multiple equilibria and payoff relevant heterogeneity together. The identification in this paper covers both static and dynamic games, while AM considers static games using the results for finite mixture models such as Kasahara and Shimotsu (2009). This paper provides detailed proofs of nonparametric identification using results in Hu (2008) for measurement error models and Hu and Shum (2012) for dynamic models with latent state variables.

3 Nonparametric Identification of Static Discrete Games

This section begins with describing static discrete games with incomplete information. Then provides the nonparametric identification methodology in both cross-sectional and panel data structures. The identification is implemented in three steps. The first step is to identify the number of equilibria. The second step is to identify strategies of each equilibrium. The last step is to identify the payoff functions.

3.1 Basic setup of static games

Consider a static simultaneous move game that involves \( N \) players. Players obtain action relevant payoff shocks before they make their decisions. These profit shocks are private information and only observable to the player herself. In each game, player \( i, i \in \{1, \ldots, N\} \), chooses an action \( a_i \) out of a finite set \( \mathcal{A} = \{0, 1, \ldots, K\} \). Let \( a_{-i} \) denote rivals’ actions and \( s \in \mathcal{S} \) denote public observable state variables. The \( K+1 \) action specific profit shocks are denoted as \( \epsilon_i(a_i) \) and their density distributions denoted as \( f(\epsilon_i) \).\(^6\) The payoff for player \( i \) when choosing action \( a_i \) is assumed to be additive separable as below:

\[
U_i(a_i, a_{-i}, s, \epsilon_i) = \pi_i(a_i, a_{-i}, s) + \epsilon_i(a_i)
\]

Unlike a standard discrete choice model, player \( i \)’s payoff not only depends on her own action but also depends on actions that her rivals choose. In particular, actions that rivals choose enter player \( i \)’s payoff function directly. This interaction among players brings in the possibility of multiple equilibria.

Assumption 1. (Conditional Independence) The random payoff shocks are identical and independent distributions (i.i.d) across actions and players, and the density distribution \( f(\epsilon_i) \) has full support and is common knowledge.

\(^6\)Similar setups are studied in Seim (2006) and Aradillas-Lopez (2010).
The assumption of conditional independence of private information is commonly imposed in the literature on estimation and inference in static games with incomplete information and social interaction models (see, e.g., Seim (2006), Aradillas-Lopez (2010), Sweeting (2009), De Paula and Tang (2012), Bajari, Hong, Krainer, and Nekipelov (2010a) as well as Ellickson and Misra (2008). This assumption can also be found in the literature on the estimation of dynamic games with incomplete information.

Given that player \( i \) only observes her own private shock \( \epsilon_i \), player \( i \)'s decision rule is a function mapping the state \( s \) and her own private information \( \epsilon_i \) to an action \( a_i \), i.e., \( a_i = \delta_i(s, \epsilon_i) \). Define \( \sigma_i(a_i|s) \) as the probability that player \( i \) chooses action \( a_i \) conditional on observing \( s \). By definition, \( \sigma_i(a_i|s) = \int I\{\delta_i(s, \epsilon_i) = a_i\}f(\epsilon_i)d\epsilon_i \)

and the expected utility of choosing action \( a_i \) by player \( i \), when she expects her rivals’ conditional choice probability (CCP) is \( \sigma_{-i}(a_{-i}|s) \), can be represented as:

\[
\begin{align*}
    u_i(a_i, s, \epsilon_i) &= \sum_{a_{-i}} \pi_i(a_i, a_{-i}, s)\sigma_{-i}(a_{-i}|s) + \epsilon_i(a_i) = \Pi_i(a_i, s) + \epsilon_i(a_i)
\end{align*}
\]

With the definition of the choice specific expected utility, the connection between the CCP of player \( i \) and the difference of the choice specific expected utility is as follows:

\[
\begin{align*}
    \sigma_i(a_i = k|s) &= F_i^k(\Pi_i(a_i = 1, s) - \Pi_i(a_i = 0, s), ..., \Pi_i(a_i = K, s) - \Pi_i(a_i = 0, s))
\end{align*}
\] (1)

Function \( F_i^k \) is determined by the joint distribution of the private information, i.e., \( f(\epsilon_i) \). As already shown in discrete choice models, it is impossible to identify both the payoff functions and the private shock distribution nonparametrically. Consequently, this paper assumes that the private shock distributions are known to the econometrician.

**Definition 1.** (BNE) For a fixed state \( s \in \mathcal{S} \), the Bayesian Nash Equilibrium (BNE) is a collection of probabilities \( \sigma_i^*(a_i = k|s) \) for \( i = 1, ..., N \) and \( k \in \mathcal{A} \) such that for all \( i \) and \( k \), equation (1) satisfied.

Similar to Hotz and Miller (1993), the equilibrium condition implies a one-to-one mapping between the CCPs of the equilibrium and the difference of expected choice utilities. Let \( \Gamma : \mathcal{A}^n \times \mathcal{S} \to [0, 1]^n \) denote the general form of the mapping, thus,

\[
\{\sigma_1^*(a_1|s), ..., \sigma_n^*(a_n|s)\} = \Gamma(\Pi_1(a_1, s) - \Pi_1(a_1 = 0, s), ..., \Pi_n(a_n, s) - \Pi_n(a_n = 0, s))
\] (2)
Assumption 1 implies that the best response probability mapping $\Gamma$ is continuously differentiable. Therefore, the mapping has at least one fixed point by Brouwer’s fixed point theorem. One can invert the mapping $\Gamma$ so that if CCPs are obtained, the differences of the expected choice values are as well. If a unique equilibrium is guaranteed, CCPs computed from collected data can be used to approximate the CCPs predicted by the exact equilibrium. As a result, the expected choice probabilities are identified. With an exclusion restriction, the payoff functions can be nonparametrically identified.

The equilibrium, however, may not be unique because the equilibrium conditions are systems of nonlinear equations. Moreover, the assumption of an unbounded error support implies that any outcome is possible in any equilibrium. Equilibria differ only in the probability assigned to individual outcomes. When multiple equilibria are presented, the one-to-one mapping does not hold anymore. The choice probabilities computed from collected data do not approximate choice probabilities of any equilibrium, instead they equal the mixture of the choice probabilities of different equilibria. Pooling choices across markets may not reflect an equilibrium anymore because the mixture of equilibria may not be an equilibrium in itself.

The approach used in current literature relies on the assumption that the same equilibrium is played across markets when the multiplicity is presented. Without any particular reason, it is not convincing that all players favor one equilibrium over others. On the other hand, without identifying the choice probabilities of any equilibrium, one cannot proceed to identify the expected choice utility, at least if one wants to follow the Hotz-Miller two-step procedure. As Jovanovic (1989) pointed out, however, multiplicity does not necessarily imply the model cannot be identified. The following section states in detail how to identify the equilibrium choice probabilities using a technique from measurement error literature.

### 3.2 Nonparametric Identification of Static Games

This section proposes an identification methodology for the static games described above allowing the presence of multiple equilibria. The identification is presented in three steps with identifying the number of equilibria, strategies and state transition of each equilibrium, and the payoff functions respectively.

Suppose the econometrician observes the actions of all players in cross-sectional markets $m$ where $m = 1, \ldots, M$ with characteristics $s_m \in \mathcal{S}$. Assume that $S$ is discrete and has finite support. Since the number of equilibria is finite (Harsanyi (1973)), I index the equilibrium by $e^* \in \Omega_{\pi,s} \equiv \{1, \ldots, Q_s\}$,
where $Q_s$ is the number of equilibria. Note that the cardinality of equilibrium set depends on the payoff and the game characteristics. Identification of equilibrium-specific components is established conditional on the market characteristics $s_m$, so I suppress the market characteristics for ease of notation. I will reintroduce market characteristics when I move to the identification of payoff functions.

With the assumption that private shocks are independent across actions and players, players’ actions are independent when the same equilibrium is selected in the cross-sectional games.

$$Pr(a_1, ..., a_n|e^*) = \prod_i Pr(a_i|e^*)$$  \hspace{1cm} (3)

However, different games might employ different equilibria, which is unobserved to econometricians. Thus, the observed joint distribution of players’ actions is a mixture over equilibrium-specific distribution, which can be represented as:

$$Pr(a_1, ..., a_n) = \sum_{e^*} Pr(a_1, ..., a_n|e^*)Pr(e^*) = \sum_{e^*} \prod_i Pr(a_i|e^*)Pr(e^*)$$  \hspace{1cm} (4)

where $Pr(e^*)$ is the probability that equilibrium $e^*$ is employed, i.e., equilibrium selection mechanism, and $Pr(a_i|e^*)$ is the CCPs of player $i$ in equilibrium $e^*$. The first equality holds because of the law of total probability. With a unique equilibrium selected in the data, even though we cannot infer which $e^*$ it is, we know that $Pr(a_1, ..., a_n)$ and $Pr(a_i)$ estimated from the data should satisfy the independent condition, i.e., $Pr(a_1, ..., a_n) = \prod_i Pr(a_i)$. As a result, this independent condition can be used to test whether the data is associated with a unique equilibrium. Failing to reject the null hypothesis that this condition holds, one can argue that all game select the same equilibrium even though the number of equilibria predicted by the model is unknown (see De Paula and Tang (2012) for a formal test).

With the presence of multiple equilibria, the correlation among players’ actions display the underlying equilibrium, which is the essential condition for identify how strategies vary in different equilibria. Before I move to the identification detail, I define identification first.

**First Step Identification**  The number of equilibria $Q$, equilibrium selection mechanism $p = \{Pr(e^*)\}_{e^*=1}^Q$ and CCPs $P = \{Pr(a_i|e^*)\}_{i=1}^n$ in different equilibria $e^*$ are identified if there does not exist another set of $\{Q, p, P\}$ that is consistent with the observed data $\{a_1^n, ..., a_n^n\}_m$.

To make full use of information for the first step identification, I divide players into two groups equally, i.e., $l = \lfloor n/2 \rfloor$ (floor of $n/2$) where $n$ is the number of players. Creating new variables for actions of players in each group, and denote the new variables as $b_1$ and $b_2$ respectively, i.e., $b_1 = \{a_1, a_2, ..., a_l\} \in \mathcal{A}^l$ and $b_2 = \{a_{l+1}, ..., a_2l\} \in \mathcal{A}^l$. Now the dimension of $\mathcal{A}^l$ becomes $(K+1)^l$. Treating
the equilibrium index as a latent variable and \( b_1 \) and \( b_2 \) as its two measurements, I can identify the number of equilibria using the rank of inequality condition described in the following. With the new variables, the joint distribution becomes:

\[
Pr(b_1, b_2) = \sum_{e^*} Pr(b_1|e^*)Pr(b_2|e^*)Pr(e^*)
\]  

(5)

I rewrite this equation into matrix form:

\[
F_{b_1, b_2} = A_{b_1|e^*}DA_{b_2|e^*}^T
\]  

(6)

where

\[
F_{b_1, b_2} \equiv [Pr(b_1 = k, b_2 = j)]_{k,j},
\]

\[
A_{b_i|e^*} \equiv [Pr(b_i = k|e^* = q)]_{k,q}
\]

\[
D \equiv \text{diag}[Pr(e^* = 1) \ldots Pr(e^* = Q)].
\]

Those matrices stack the distributions with all possible values that \( b_1 \), \( b_2 \) and \( e^* \) can take. In particular, matrix \( F_{b_1, b_2} \) consists of the whole joint distributions between \( b_1 \) and \( b_2 \), which can be estimated from data. \( D \) is a diagonal matrix with the probability of each equilibrium being selected as diagonal elements, while matrix \( A_{b_i|e^*} \) collects all the CCPs in every equilibrium. If \( n = 2, K = 1 \) and \( Q = 2 \), matrix \( A_{b_i|e^*} \) becomes:

\[
A_{a_i|e^*} = \begin{pmatrix}
Pr(a_i = 0|e^* = 1) & Pr(a_i = 0|e^* = 2) \\
Pr(a_i = 1|e^* = 1) & Pr(a_i = 1|e^* = 2)
\end{pmatrix}
\]

The dimensions of the three matrices defined above \( F_{b_1, b_2} \), \( A_{b_i|e^*} \) and \( D \) are \((K + 1)^l \times (K + 1)^l\), \((K + 1)^l \times Q\), and \( Q \times Q \) respectively. Note that the number of equilibria \( Q \) is unknown. As I will show, the number of equilibria \( Q \) is identifiable from data under further assumptions. This contrasts with the existing literature, which often assumes a unique equilibrium. The identification of the number of equilibria is summarized in the following lemma.

**Lemma 1.** The rank of the observed matrix \( F_{b_1, b_2} \) serves as the lower bound for the number of equilibria, i.e., \( Q \geq \text{Rank}(F_{b_1, b_2}) \). Furthermore, the number of equilibria is identified, particularly, \( Q = \text{Rank}(F_{b_1, b_2}) \) if the following conditions are satisfied:

(1) \((K + 1)^l > Q\); (2) both matrices \( A_{b_1|e^*} \) and \( A_{b_2|e^*} \) are full rank; (3) all \( Pr(e^*) \) are positive.

**Proof** See Appendix B. □

10
The first condition requires that the number of possible action combination of the two groups is greater than the number of equilibria. Note that the action variable serves as a measurement for the latent variable. Thus, sufficient variation is needed to infer the dimension of the underlying equilibrium.

The full rank condition implies that CCPs in any equilibria are not a linear combination of CCPs in any other equilibria. This condition essentially requires that no equilibrium is redundant. Identification power comes from the fact that players respond to different equilibria via choosing alternative actions differently. Note that \( b_1 \) and \( b_2 \) are divided arbitrarily, while I can compute the rank of matrix \( F_{b_1, b_2} \) for different division of players and keep the division which generates the highest rank. As a result, the full rank condition is not as restrictive as it sounds because the identification only requires the full rank of matrices \( A_{b_1|e^*} \) and \( A_{b_2|e^*} \) for one division. As long as there exists one way to partition the players so that full rank condition holds, the number of equilibria is identified. Moreover, when the number of equilibria equals two, the full rank condition holds naturally because CCPs are different in both equilibria for at least two players.

The third condition indicates that only those equilibria that are active can be identified. This also means that we might not be able to recover how many equilibria are actually predicted by the model. However, it is not easy to compute all the equilibria even using the homotopy method, which by itself is computationally challenging and time consuming. Furthermore, it is not necessary to recover all the equilibria during estimation of game primitives. Most importantly, this condition does not affect estimation of game primitives.

Identification of the number of equilibria is empirically important. The theoretical model does not provide much guidance as to how many equilibria exactly exist in the static game of incomplete information. Bajari, Hong, Krainer, and Nekipelov (2010a) use the Homotopy method to compute all the equilibria in a special static game of incomplete information. They find that the number of equilibria decreases with the number of players and the number of possible individual alternatives that they are allowed to choose from. Lemma 1 provides a plausible approach to determine \( Q \) from the data.

Provided that \( Q \) is identified, I show below how to identify CCPs of different equilibria for each individual player. Without loss of generality, suppose data from two players is sufficient to identify the number of equilibria, i.e., \( K + 1 > \text{rank}(F_{b_1, b_2}) = Q \). In measurement error literature, normally the latent variable share the same support with its measurements, and the full rank condition means those matrices are invertible. As a result, identification utilizes the matrix eigenvalue-eigenvector decomposition technique. Following the same technology, I partition the dimension of space for measurements.
to the equilibrium index so that they have the same support, i.e., $A = \{1, 2, \ldots, Q\}$. There exists a function $g$ that $g : \mathcal{A} \rightarrow \tilde{\mathcal{A}}$ and $\tilde{a}_i = g(a_i)$. The function $g$ is created in a way the matrix $A_{\tilde{a}_i|e^*}$, which is defined in the same way as $A_{a_i|e^*}$, is full rank for both $i = 1, 2$. The following lemma states that $g$ always exists.

**Lemma 2.** For a matrix $F$ with dimension of $K + 1 \times K + 1$ and rank is $Q$, there exists a way to partition it into a $Q \times Q$ matrix with rank of $Q$.

**Proof** See Appendix B

Since $\tilde{a}_i$ is a function of $a_i$, and $a_i$ are independent with each other, $A_{\tilde{a}_i|e^*}$ is also independent between each other. Identification requires extra information, i.e., player 3. Given that the conditional independence condition holds for any number of players, I can obtain the following joint distribution expression.

$$Pr(\tilde{a}_1, \tilde{a}_2, a_3) = \sum_{e^*} Pr(\tilde{a}_1|e^*)Pr(\tilde{a}_2|e^*)Pr(a_3|e^*)Pr(e^*)$$

(7)

Now I introduce more matrix notations:

$$F_{\tilde{a}_1, \tilde{a}_2, a_3 = k} = [Pr(\tilde{a}_1 = l, \tilde{a}_2 = j, a_3 = k)]_{lj},$$

$$A_{\tilde{a}_i|e^*} = [Pr(\tilde{a}_i = l|e^* = q)]_{lj},$$

$$F_{\tilde{a}_1, \tilde{a}_2} = [Pr(\tilde{a}_1 = l, \tilde{a}_2 = j)]_{lj},$$

$$D_{a_3 = k|e^*} = \text{diag}
\left[Pr(a_3 = k|e^* = 1) \cdots Pr(a_3 = k|e^* = Q)\right].$$

Using the above matrix representations, we have following two equations:

$$F_{\tilde{a}_1, \tilde{a}_2} = A_{\tilde{a}_1|e^*}DA_{\tilde{a}_2|e^*}^T$$

(8)

$$F_{\tilde{a}_1, \tilde{a}_2, a_3 = k} = A_{\tilde{a}_1|e^*}D_{a_3 = k|e^*}DA_{\tilde{a}_2|e^*}^T$$

(9)

Given that $A_{\tilde{a}_1|e^*}$ and $A_{\tilde{a}_2|e^*}$ have full rank, post-multiplying $F_{\tilde{a}_1, \tilde{a}_2}^{-1}$ on both sides of equation 9 leads to the following main equation, which is essential for the identification.

$$F_{\tilde{a}_1, \tilde{a}_2, a_3 = k}F_{\tilde{a}_1, \tilde{a}_2}^{-1} = A_{\tilde{a}_1|e^*}DA_{a_3 = k|e^*}A_{\tilde{a}_1|e^*}^{-1}$$

(10)

The right-hand side of the equation above represents an eigenvalue-eigenvector decomposition of the matrix on the left-hand side, with $D_{a_3 = k|e^*}$ being the diagonal matrix consisted of eigenvalues, and $A_{\tilde{a}_1|e^*}$
being the eigenvector matrix, see Hu (2008). The left-hand side of the equation can be estimated from the observed data, therefore this equation can be used to identify both $D_{a_3=k|e^*}$ and $A_{\tilde{a}_1|e^*}$ simultaneously.

To fully identify the model, uniqueness of the decomposition is required. Namely, eigenvalues are distinctive. I impose the following assumption on CCPs of different equilibria to guarantee a unique decomposition:

**Assumption 2. (Distinctive Eigenvalues)** There exists an action $k$ of player 3, such that for any two equilibria $i \neq j$, the probability of this action taken under different equilibria is different, i.e., $Pr(a_3 = k|e^* = i) \neq Pr(a_3 = k|e^* = j)$.

This assumption rules out the possibility that choice probabilities of a typical action for a player from different equilibria are the same. The distinctive assumption is empirically testable because matrix $F_{\tilde{a}_1, \tilde{a}_2, a_3=k}F^{-1}_{\tilde{a}_1, \tilde{a}_2}$ can be estimated from the data. Note that I do not require the eigenvalues of $F_{\tilde{a}_1, \tilde{a}_2, a_3=k}F^{-1}_{\tilde{a}_1, \tilde{a}_2}$ to be distinct for every $a_3 = k$. As long as there exists one $a_3 = k$ such that distinct eigenvalues are guaranteed, the analysis is valid.

Upon using assumption 2, the eigenvalue-eigenvector decomposition in equation 10 identifies $A_{\tilde{a}_1|e^*}$ and $D_{a_3=k|e^*}$ up to a normalization and ordering of the columns of the eigenvector matrix $A_{\tilde{a}_1|e^*}$. Note that each column of the eigenvector matrix $A_{\tilde{a}_1|e^*}$ is a whole conditional distribution for one equilibrium, hence the column sum of the matrix equals one. This column sum property can be used for normalization of the eigenvector matrix. Since the index of equilibria does not include meaning for identifying the original model, all we need to know is how many equilibria are built in the model and the CCPs under each equilibrium. Eigenvalues are not required to be any specific ordering. Thus, no extra assumption is needed for ordering of eigenvalues. Any ordering is fine.

With the decomposition, I can identify CCPs of player 3 for a typical value of $a_3$. Identification of CCPs for other actions and other players are provided in Appendix B. Consequently, CCPs of players in each equilibrium are identified, which is stated in the following lemma.

**Lemma 3.** With assumptions 1 and 2, and the conditions in lemma 1 satisfied, CCPs of players in each equilibrium and the equilibrium selection are nonparametrically identified.

**Proof** See Appendix B

Based on lemmas 1, 2, and 3, conditional on market characteristic $s$, I have already identified all the
CCPs under every employed equilibrium $Pr(a_i|s,e^* = 1) .... Pr(a_i|s,e^* = q)$ should satisfy the original equilibrium condition. From the discrete choice literature, it is not possible to identify both the mean utility functions and the joint distribution of the error terms without making strong exclusion and identification at infinity assumptions (see for example Matzkin (1992)). Here I assume the distributions of private shocks are known. More specifically, I assume the error terms follow extreme value distribution, upon which the equilibrium condition becomes

\[ \log \sigma_i(a_i = k|s,e^*) - \log \sigma_i(a_i = 0|s,e^*) = \sum_{a_{-i}} \left( \pi_i(a_i = k,a_{-i},s) - \pi_i(a_i = 0,a_{-i},s) \right) \sigma_{-i}(a_{-i}|s,e^*) \]

As with analysis in discrete choice models, it is impossible to separately identify all the payoff functions, but only their differences $\pi_i(a_i = k,a_{-i},s) - \pi_i(a_i = 0,a_{-i},s)$. Normalization is necessary and stated in the following:

**Assumption 3. (Normalization)** For all $i$ and all $a_{-i}$ and $s$, $\pi_i(a_i = 0,a_{-i},s) = 0$.

This assumption sets the mean utility from a particular choice equal to zero, which is similar to the outside good assumption in the discrete choice model. If we aim at looking into how firms strategically interact with each other, i.e., how one’s actions affect profits of others, this normalization does not affect our analysis. With the normalization condition, the equilibrium condition becomes

\[ \log \sigma_i(a_i = k|s,e^*) - \log \sigma_i(a_i = 1|s,e^*) = \sum_{a_{-i}} \pi_i(a_i = k,a_{-i},s) \sigma_{-i}(a_{-i}|s,e^*) \]

Identification of payoff functions $\pi_i(a_i = k,a_{-i},s)$ requires exclusion restrictions (see Bajari, Hong, Krainer, and Nekipelov (2010b) and Bajari, Hahn, Hong, and Ridder (2011)). If there are covariates that shift the utility of one player, but can be excluded from the utility of other players, then all the payoffs can be identified nonparametrically. As a result, to identify the payoff functions, I state the exclusion restriction assumption below:

**Assumption 4. (Exclusion Restriction)** For each player $i$, the state variable can be partitioned into two parts denoted as $s_i,s_{-i}$, so that only $s_i$ enters player $i$’s payoff, i.e. $\pi_i(a_i = k,a_{-i},s) = \pi_i(a_i = k,a_{-i},s_i)$.

An example of exclusion restrictions is a covariate that shifts the profitability of one firm but that can be excluded from the profits of all other firms. Firm specific cost shifters are commonly used in empirical work. For example, Jia (2008) and Holmes (2011) demonstrate that distance from firm
headquarters or distribution centers is a cost shifter for big box retailers such as Walmart. With the exclusion restriction, the above equation becomes

\[
\log \sigma_i(a_i = k | s_i, s_{-i}, e^* = q) - \log \sigma_i(a_i = 0 | s_i, s_{-i}, e^* = q) = \sum_{a_{-i}} \pi_i(a_i = k, a_{-i}, s_i) \sigma_{-i}(a_{-i} | s_i, s_{-i}, e^* = q)
\]

Variation of \(s_{-i}\) expands the total number of equations without adding more unknowns, which helps for identifying payoff functions \(\pi_i(a_i = k, a_{-i}, s_i)\) nonparametrically.

With the idea of exclusion restrictions, one can see that the existence of multiple equilibria helps for the identification of payoff functions. Equilibrium shifts the choice probabilities without shifting the payoff functions. Essentially, it plays a role as an exclusion restriction. However, enough variation of the exclusion restriction is needed to identify the payoff functions nonparametrically. There is not enough variation from multiple equilibria because the number of equilibria is relatively small compared to the number of actions or players for the first step identification.

Specifically, fixing \(s\), there are \(K \times Q_s\) equations, which is magnified by the number of equilibria \(Q_s\), while there are \(K \times (K+1)^{n-1}\) unknowns. To nonparametrically identify the profit function requires that the number of equations is greater than the number of unknowns \((Q_s \geq (K+1)^{n-1})\). Unfortunately, first step identification requires the number of multiple equilibria \(Q_s\) to be smaller relative to the number of actions \((K+1)\) or the number of players \((n)\). These two conditions conflict with each other. Consequently, with multiple equilibria itself as an exclusion restriction, I cannot nonparametrically identify payoff functions. Even though presence of multiple equilibria does not enable one to identify the payoff functions, it lessens the burden of the variation for extra exclusion restrictions. The importance of multiple equilibria is shown when variation from available exclusion restrictions is not enough, or there is no exclusion restriction at all (see Sweeting (2009)).

**Theorem 1.** (Identification of static game) With assumptions 1, 2, 3, and 4 and conditions in lemma 1, static games with incomplete information can be nonparametrically identified. Specifically, the number of equilibria, the equilibrium selection, the CCPs of player in each equilibrium and the payoff functions are all nonparametrically identified.

### 3.3 Nonparametric Identification with Panel Data in Static Games

When only cross-sectional data is available, estimation can only be done by pooling data from different games. With panel data, estimation can be done along time series on individual markets, which im-
explicitly assumes that the same equilibrium is employed over time in individual markets. Unfortunately, there is hardly any theoretical or empirical evidence to support this assumption. Even if it is true that players employ the same equilibrium over time, another consideration is that estimation requires enough observation for the same game. Moreover, without knowing the underlying equilibrium, it is impossible for policy implication. Thus, this subsection presents identification of static games with panel data relaxing the same equilibrium assumption.

Denote \( a_{it}, i = 1, ..., n \) as the action chosen by player \( i \) in period \( t \). In each period, pool the actions of all players together and denote it as \( a_t \), i.e., \( a_t = \left( a_{1t}, ..., a_{nt} \right) \in \mathcal{A} \) where \( \mathcal{A} \) has a dimension of \( M = (K + 1)^n \). Note that identification power comes from the variation of measurements of the latent variable. Taking all players’ action as a whole allows me to identify a larger number of equilibria than in the case of cross-sectional data. Let \( e_{it}^* \) denote the index of equilibria employed in time \( t \), I assume that \( e_{it}^* \) follows a first-order Markov process. This paper is not going to model how the players choose different equilibria over time.

**Assumption 5.** (First-order Markov Equilibrium Evolution) The equilibrium that a typical market employs follows a first-order Markov process, i.e., \( \Pr(e_{it+1}^* | e_{it}^*, ..., e_{i0}^*) = \Pr(e_{it+1}^* | e_{it}^*) \).

With this assumption, the correlation between actions in different periods comes from the evolution of underlying equilibria. Thus, actions in different periods are independent conditional on underlying equilibria, which is the key for identification. This assumption nests the conventional assumption of the same equilibrium by an identity transition matrix. Allowing higher order of Markov process is possible given a longer period of data. According to the law of total probability, the observed joint distribution for actions in two periods can be represented as:

\[
\Pr(a_{t+1}, a_t) = \sum_{e_{t+1}^*} \Pr(a_{t+1}, e_{t+1}^*, a_t) = \Pr(a_{t+1} | e_{t+1}^*) \Pr(e_{t+1}^*) \Pr(e_{t+1}^*)
\]

where \( \Pr(a_t | e_{t+1}^*) \) represents the probability of the players choosing action \( a_t \) in period \( l \) when the equilibrium chosen in period \( t \) is \( e_{t+1}^* \); \( \Pr(e_{t+1}^*) \) is the fraction of markets that employ equilibrium \( e_{t+1}^* \) at period \( t + 1 \), i.e., the marginal distribution of the equilibrium index in period \( t + 1 \). Note that private information is allowed to be correlated across different players because the identification relies on the variation across time.

Similar to the logic of the cross-sectional case, identification of the number of equilibria is through the rank of a matrix constructed of the joint distribution of \( a_t \) and \( a_{t+1} \). The intuition is that \( a_t \)
and $a_{t+1}$ are correlated through the underlying equilibrium $e_{t+1}^*$, otherwise they are independent, (see Appendix for detail). One advantage of using panel data is that the support of the measurements is bigger. Identification is feasible for more equilibria if we have more periods of data because we can pool actions from different periods together to expand our choice set to help identify the case in which the number of equilibria is bigger.

With the number of equilibria identified, I next proceed to identify CCPs. This requires an extra period of data. From the observed joint distribution,

$$Pr(a_{t+2}, a_{t+1}, a_t) = \sum_{e_{t+1}^*} Pr(a_{t+2}|e_{t+1}^*) Pr(a_{t+1}|e_{t+1}^*) Pr(a_t|e_{t+1}^*) Pr(e_{t+1}^*)$$

Sum over $a_{t+1}$ leading to:

$$Pr(a_{t+2}, a_t) = \sum_{e_{t+1}^*} Pr(a_{t+2}|e_{t+1}^*) Pr(a_t|e_{t+1}^*) Pr(e_{t+1}^*)$$

Identification of CCPs and the Markov evolution relies on the above two equations connecting the observed and unknowns and the identified number of equilibria. First of all, I map action space into a new space with a support of $Q$. Then CCPs in each equilibrium can be identified as eigenvalues of the observed matrices consisting of joint distributions (see Appendix for detail). Lastly, payoff functions can be nonparametrically identified with exclusion restrictions, which is exactly the same as in the cross-sectional case. As a result, I summarize the identification results with panel data in the following theorem:

**Theorem 2.** *(Identification of Static Game with Panel Data)* Under assumptions 3, 4, and 5, and an assumption of distinctive eigenvalues, the number of equilibria $Q$, CCPs of players in each equilibrium, the equilibrium evolution and payoff functions are nonparametrically identified using three periods of data.

Like static games, multiple equilibria possibly occurs in dynamic games. Moreover, dynamic games are a generalization of the static frameworks with future discount factor being zero. The presence of multiple equilibria in dynamic games also results in a mixture feature. As a result, I can follow a similar approach to identify dynamic games, which is presented in the next section.
4 Semiparametric Estimation

The proposed identification methodology is constructive and simple, so estimation can also be achieved in two steps by following the identification procedure. Instead of purely nonparametric estimation, this section provides a nonparametric estimation of the number of equilibria and CCPs, but a parametrical estimation of payoff functions. Note that although this section only provides estimators for static games, the analog counterpart can be obtained for dynamic games.

4.1 Nonparametric Estimation of The Number of Equilibria and CCPs

To estimate the number of equilibria from the data directly, one needs to estimate the joint distribution of actions by players. With the assumption of discrete states, the joint distribution can be estimated via simple frequency, regardless of multiplicity. Simple frequency estimators are not feasible when states are continuous. However, if a unique equilibrium is guaranteed for all states, sieve series expansions can be used to estimate the joint distribution\(^7\). This is because the joint distribution is continuous along the state variables. Also, other nonparametric regression methods such as kernel smoothing or local polynomial regressions can be used to obtain the joint distribution.

However, those conventional estimation approaches are invalid when multiple equilibria are present. This is because the joint distribution is no longer continuous along the state variables. We have to use estimation methods that can deal with this potential discontinuity problem. Muller (1992) provides a methodology to detect the discontinuous point and estimate corresponding discontinuous functions. One restriction of this approach is that it can only detect a finite number of discontinuous points.

Additional assumptions are needed, such as continuity of the equilibrium selection mechanism, if we want to apply the Muller (1992) approach here. First of all, the number of equilibria is not continuous by nature. It will jump along the state variable, resulting in a finite number of discontinuous points. Moreover, for those state variables that share the same number of equilibria, the equilibrium selection mechanism needs to be continuous with finite discontinuous points so that the joint distribution of observed actions is continuous due to the mixture feature. To sum up, when state variables are continuous and multiple equilibria are present, to estimate the conditional joint distributions of observed actions, the equilibrium selection mechanism has to be continuous almost everywhere with a finite discontinuous points.

\(^7\)See Newey (1990), Ai and Chen (2003) and Newey (1994) for how to use sieve series expansions to estimate.
number of discontinuous points.

This paper assumes that state variables are discrete, so the joint distributions of players can be estimated through the simple frequency:

$$\hat{P}_r(a_1, \ldots, a_n|s) = \frac{1}{M} \sum_{m} I(a_1^m = a_1, \ldots, a_n^m = a_n, s^m = s)$$

With estimation of the joint distributions, matrix $F_{a_1,a_2}$ is obtained since it essentially collects joint distributions of actions by player 1 and 2. Since the rank of matrix $F_{a_1,a_2}$ is used to infer the number of equilibria, one must estimate the rank of a matrix from the matrix’s estimator.

There is a lot of ongoing research on testing rank of a matrix through the estimates of that matrix, such as the singular value decomposition in Kleibergen and Paap (2006) and characteristic roots of a quadratic form built from the matrix in Robin and Smith (2000). See also Camba-Mendez and Kapetanios (2009) for a review. Those testing methodologies are proposed for general matrices. The matrices used here are square, and we eventually need to partition the actions so that the joint distribution matrix defined by new partition options is full rank. The full rank condition for a square matrix is equivalent to the fact that the determinant of this matrix is positive. Thus the determinant is a natural statistic to be used for testing in this paper.

A sequence of tests is performed to estimate the rank of the matrix constructed by those joint distributions. Suppose there are $n$ players in the game, denote $m = \lfloor n/2 \rfloor$, $s_1 = \{a_1, \ldots, a_m\}$ and $s_1 = \{a_{m+1}, \ldots, a_{2m}\}$, thus the matrix $F_{s_1,s_2}$ with dimension of $(K+1)^m \times (K+1)^m$ can be estimated through $\hat{F}_{s_1,s_2}$. Let $\hat{F}_{s_1,s_2}^J$ denote the matrix with dimension $J \times J$ through partitioning of the original $s_1$ and $s_2$ into $J$ number of support. The sequential procedure is given as follows: the sequence of hypotheses $H_0^J : det(F^J) = 0$ is tested against the alternatives $H_1^J : det(F^J) \neq 0$ in decreasing order starting at $J = (K+1)^m$. If no rejection occurs until $J = 1$, set $Q = 1$. Each hypothesis in this sequence is tested by a t-test where the error variance is always estimated from the overall model. More formally, we have,

$$\hat{r} = \min\{J : |T_J| \geq c_jN, 0 \leq J \leq (K+1)^m\}$$

where $N$ represents the sample size. If the critical value associated with the significant level $\alpha$ is set to be a constant, then the rank estimated through this testing procedure is not consistent. The reason is that the testing procedure rejects the true with probability $\alpha$ even as the sample size goes to infinity. To obtain consistent estimates for the rank, the critical value is chosen associated with a sample dependent
significant level in a way that $\alpha_N$ goes to zero as the sample size $N$ goes to infinity but not faster than a given rate. Hosoya (1989) shows that if $\alpha_N$ goes to zero as the sample size $N$ goes to infinity and also $\lim_{N \to \infty} \frac{\ln \alpha_N}{N} = 0$, then the rank estimate provided by the sequential testing procedure will converge in probability to the true rank.

Note that in order to estimate the number of equilibria, one additional condition needs to be satisfied. That is, the dimension of the matrix used to infer the number of equilibria is greater than the number of equilibria. As a result, if one cannot reject the null in the very first step, then only the lower bound of the number of equilibria is obtained. One cannot tell the exact number of equilibria from the data in hand. Consequently, both point identification and estimation can not be obtained. However, one can always turn to set identification, which might provide useful information for inference.

With the number of equilibria consistently estimated, the estimation of the CCPs under different equilibria exactly follows the identification procedure. For example, assume that the number of equilibria is less than the action choices of one individual player, and information from three players is sufficient to estimate all the CCPs. For a typical action from player 3 that satisfies the distinct eigenvalue condition, I first do the decomposition for this typical actions and the CCPs of player 3 choosing this action can be obtained as eigenvalues through decomposition. Then all other CCPs can be estimated. Additionally, the equilibrium selection mechanism can be estimated through equation 10.

4.2 Parametric Estimation of Payoff Function

With CCPs under each equilibrium estimated, payoff functions can be estimated nonparametrically with exclusion restrictions by following the identification procedure. Here I parameterize the payoff function and estimate the structural parameters using a prevalent two-step estimation method. Denote the parameterized payoff functions as $\pi_i(a_i, a_{-i}, s) = \pi_i(a_i, a_{-i}, s; \theta)$, and suppose market characteristics are discrete with dimension of $d$, $s \in \mathcal{S} = \{s_1, ..., s_d\}$.

Pioneered by Hotz and Miller (1993)\textsuperscript{8}, two-step estimators are widely used for estimation in discrete choice models, static and dynamic games. Comparing to the Nested Fixed Point Theorem algorithm by Rust (1987), two-step estimators are computationally light because they do not need to solve for

\textsuperscript{8}For other two-step estimators, see the pseudo-maximum likelihood estimator by Aguirregabiria and Mira (2002), and estimators for dynamic games recently considered in Aguirregabiria and Mira (2007), Pakes, Ostrovsky, and Berry (2007), and in Bajari, Benkard, and Levin (2007). See also Pesendorfer and Schmidt-Dengler (2008) for a unified framework of two-step estimators.
the fixed point. It is well known that looking for a fixed point is computationally challenging and time consuming. Two-step estimators begin with consistently estimating the auxiliary choice probabilities in the first step, and then recovering the structural parameters through constraints from equilibrium conditions. As a result, in order to obtain well-behaved estimators for the structural parameters, the auxiliary choice probabilities need to be estimated consistently at the beginning. Otherwise, the error will be augmented and the second step estimator will behave poorly. This is why in previous literature the existence of multiple equilibria makes two-step estimators invalid. The choice probabilities estimated directly from the data directly do not come from any equilibrium anymore. Instead, it is a mixture of the equilibria, which itself is not an equilibrium.

The methodology above allows me to use a two-step estimator even in the presence of multiple equilibria. Denote the first step estimates as $\hat{\sigma}(a|s,e^*)$. The equilibrium condition is represented by a general mapping denoted as $h(\sigma, \theta) = \sigma(a|s,\theta) - \Gamma(\sigma(a|s,\theta);\theta) = 0$, which holds for every $s$. The least squares estimator estimates the parameters of interest by forcing the constraints:

$$h(\hat{\sigma}, \theta) = \hat{\sigma}(a|s,e^*) - \Gamma(\hat{\sigma}(a|s,e^*);\theta) = 0$$

satisfied approximately for every $s$ and every equilibrium. With the number of equations greater than the number of parameters, a weight is assigned to individual equations for minimization. Denote $\hat{\sigma}_M$ as the vector of collecting all $\hat{\sigma}(a|s,e^*)$ and $\Gamma(\hat{\sigma};\theta)$ as another vector collects all $\Gamma(\hat{\sigma}(a|s,e^*);\theta)$. Let $W_M$ be a symmetric positive definite matrix with dimension of $((K+1)^n \cdot \sum Q_s) \times ((K+1)^n \cdot \sum Q_s)$ that may depend on the observations. A least square estimator associated with weight matrix $W_M$ is a solution $\hat{\theta}(W_M)$ to the problem

$$\hat{\theta}(W_M) = \text{argmin}_\theta \ [\hat{\sigma} - \Gamma(\hat{\sigma};\theta)]^t W_M [\hat{\sigma} - \Gamma(\hat{\sigma};\theta)]$$

Thus, the asymptotic least squares estimator $\hat{\theta}(W_M)$ brings the constraint closest to zero in the metric associated with the scalar product defined by $W_M$. A simple example of the weight matrix $W_M$ is the identity matrix, which treats all constraints equally. Another example of the weighting matrix is to weight each market type differently, according to the number of observations each type has. With regular assumptions on the payoff function, such as continuity and the first step estimators are consistent and asymptotically distributed, the structural parameters estimated through least squares are consistent and asymptotically distributed. Assumptions and proofs of asymptotical properties of the estimators are included in the Appendix C.
5 Nonparametric Identification in Dynamic Games

This section begins with describing the dynamic game framework with incomplete information, followed by the equilibrium characterization. Then I show the identification of the games allowing multiple equilibria.

5.1 Models of Dynamic Games

Consider a model of discrete time, infinite-horizon games with \( N \) players\(^9\). At the beginning of each period \( t \in \{0, 1, ..., \infty\} \), player \( i \in \{1, ..., N\} \) receives her own private profit shock \( \epsilon_{it} \) before choosing her action out of a finite set. Let \( \epsilon_t \) represent the private information for all players, i.e. \( \epsilon_t \equiv (\epsilon_{1t}, ..., \epsilon_{Nt}) \). Player \( i \)'s action is denoted as \( a_{it} \), where \( a_{it} \in \mathcal{A} = \{0, 1, ..., K\} \). Let \( a_t \) denote the action vector for all players in period \( t \), i.e., \( a_t = \{a_{1t}, ..., a_{Nt}\} \in \mathcal{A}^N \). Denote the game and players’ attributes in period \( t \) as \( x_t \in \mathcal{X} \), where \( \mathcal{X} \) has a finite support. Assume that state \( x_t \) evolves according to a stationary transition function denoted as \( g(x_t | x_{t-1}, a_{t-1}) \), which is common knowledge among all the players.

In empirical applications, actions chosen in the previous period play a role in players’ current decisions. For example, in Sweeting (2011), the format which music stations choose to air in the current period depends on the format they aired last period because of the switching cost. Furthermore, in dynamic oligopoly frameworks, firms are allowed to enter or exit the market, under which players’ current decision depends on what they chose in the last period. Moreover, to provide a more general framework for identification, here I also allow the past actions affect players’ current period payoff. As a result, past period’s action enters the policy function.

As in the dynamic games literature (Ericson and Pakes (1995)), I only consider Markov Perfect Equilibria (MPE), in which each player’s strategies can be conditioned only on the current state of the game, and the state only contains payoff relevant information. Thus, the state variable in period \( t \) for MPE is captured as \( s_t = \{x_t, a_{t-1}\} \in \mathcal{X} \times \mathcal{A}^N \), which is also discrete. The payoff for player \( i \) from choosing action \( a_{it} \) while her rivals choosing actions \( a_{-it} \) in period \( t \) is defined as follows:

\[
u_i(a_t, s_t, \epsilon_{it}) = \pi_i(a_{it}, a_{-it}, s_t) + \epsilon_i(a_{it})\]

It is standard in the literature to assume the private shocks to enter the payoff function separable.\(^9\)

additively. Given that the game is stationary, I suppress the \( t \) for ease of notation. Each period, player \( i \)'s policy function is denoted as \( \delta_i(s, a', \epsilon_i) \) with a corresponding CCP denoted as \( \sigma_i(s, \epsilon_i) \). Each period, player \( i \)'s problem is to maximize her own lifetime expected utility discounting by \( \beta \). Let \( W_i(s, \epsilon_i; \sigma) \) be player \( i \)'s value function given both public state \( s \) and her own private information \( \epsilon_i \), thus

\[
W_i(s, \epsilon_i; \sigma) = \max_{a_i \in A_i} \{ \Pi_i(a_i, s) + \epsilon_i(a_i) + \beta \int \sum_{a_{-i}} W_i(s', \epsilon_i'; \sigma) g(s'|s, a_i, a_{-i}) \sigma_{-i}(a_{-i}|s) f(\epsilon_i') d\epsilon_i' ds' \}
\]

where \( \Pi_i(a_i, s) = \sum_{a_{-i}} \pi_i(a_i, a_{-i}, s) \sigma_{-i}(a_{-i}|s) \). The first term is the certain part of the current period’s payoff, and the latter one captures future lifetime utility with discounting by \( \beta \). To define the Markov Perfect Equilibrium using choice probabilities \( \{\sigma_i(a_i|s)\}_i \) instead of the original policy functions \( \delta_i(s, \epsilon) \), first I define the choice specific value function as

\[
V_i(a_i, s) = \Pi_i(a_i, s) + \beta EW_i(s', \epsilon_i'; \sigma)
\]

Similar to the choice specific utility in static games, I define the choice specific value function as the certain part of the utility from choosing that action, excluding the additive profit shocks \( \epsilon_i(a_i) \). With representation of the choice specific value function, given \( s \) and \( \epsilon_i \), player \( i \) will choose the action that gives her the highest utility. Thus, the CCP \( \sigma_i(a_i, s) \) can be computed as:

\[
\sigma_i(a_i = k, s) = \Pr (V_i(a_i = k, s) + \epsilon_i(a_i = k) \geq V_i(a_i = j, s) + \epsilon_i(a_i = j), \forall j)
\]

Notice that for any set of strategies, the choice specific value functions, value functions and current expected utilities depend on players’ strategies only through the choice probabilities \( \sigma \) associated with decision rule \( \delta \). Following Milgrom and Weber (1985), Markov Perfect Equilibrium is defined in terms of CCP \( \sigma_i(a_i, s) \) in probability space as follows.

**Definition 2.** (MPE) A Markov Perfect Equilibrium (MPE) is a collection of CCPs \( \{\sigma_i(a_i, s)\}_i \) such that for all player \( i \) and \( s \), the following conditions are satisfied:

\[
\sigma_i(a_i = k, s) = \Pr (V_i(a_i = k, s) + \epsilon_i(a_i = k) \geq V_i(a_i = j, s) + \epsilon_i(a_i = j), \forall j)
\]

where

\[
V_i(a_i, s) = \Pi_i(a_i, s) + \beta EW_i(s', \epsilon_i'; \sigma)
\]

\[
W_i(s, \epsilon_i; \sigma) = \max_{a_i \in A_i} \{ V_i(a_i, s) + \epsilon_i(a_i) \}
\]

MPE means that players actions are fully determined by the current vector of state variables. Intuitively, whenever a player observes the same vector of states it will take the same actions and the
history of the game until period \( t \) does not influence players decisions. As in the static framework, the private information is assumed to be extreme value distribution, and the equilibrium probabilities and the choice specific value functions are related through the following equation for all \( i \) and \( k \).

\[
\sigma_i(a_i = k, s) = \frac{\exp(V_i(a_i = k, s))}{\sum_j \exp(V_i(a_i = j, s))}
\]

Again, the equilibrium conditions provide us a system of nonlinear equations. As a result, multiple equilibria are possible. Considering the multiplicity explicitly, the next subsection proposes identification of dynamic games with incomplete information.

### 5.2 Nonparametric Identification of Dynamic Games

Similar to static games, the presence of multiple equilibria is a feature inherent to dynamic games\(^\text{10}\). As discussed in Pesendorfer and Schmidt-Dengler (2008), the Markovian assumption implies that a single equilibrium is played in a market-level time series. Consequently, identification and estimation can be obtained using a single path of play to get around the multiplicity concerns. Using information from a single path, however, has other disadvantages. First of all, leaving out information of cross-sectional markets reduces efficiency of the estimation. Secondly, relying on information from an individual market over time requires a long period of data, limiting the application scope. Finally, without information of equilibrium selection, it is difficult to conduct counterfactual analysis. As a result, this section presents identification of dynamic games by making full use of cross-sectional data, allowing the existence of multiple equilibria.

Suppose there are \( N \) players playing an infinite horizon dynamic game in markets \( m = 1, 2, \ldots, M \). Data includes players’ action vector \( a_t \in \mathcal{A}^N \) and market and players characteristics \( x_t \in \mathcal{X} \) in each period, i.e., \( \{(a_t^m, x_t^m), \ldots, (a_T^m, x_T^m)\}^m \). The time-series data for each individual market is associated with the same equilibrium, while different paths of data might associated with different equilibria. Consequently, the correlation between observables along the time path reflects the aggregation of the correlation from different equilibria. As shown in Haller and Lagunoff (2000), stochastic dynamic games also have a finite number of equilibria, so I can index the equilibrium as \( e^* \in \Omega \equiv \{1, 2, ..., Q\} \), where \( Q \) is the number of equilibria.

To fully use variation from observables, I create a new variable that includes all observables in a

---

\(^{10}\)Pesendorfer and Takahashi (2012) propose several statistical tests to examine multiplicity of equilibria in a similar setup.
period, i.e., $w_t = \{a_t, x_t\} \in \mathcal{W} \equiv \mathcal{A}^N \times \mathcal{X}$. The dimension of $\mathcal{W}$ is denoted as $H$. With the state transition assumption $g(x_t | x_{t-1}, a_{t-1})$ and the fact that the data comes from MPE, $w_t$ follows a stationary first-order Markov process when the same equilibrium is selected.

$$
Pr(w_t | w_{t-1}, ..., w_1, e^*) = Pr(a_t, x_t | a_{t-1}, x_{t-1}, ..., a_0, x_0, e^*)
$$

$$
= Pr(x_t | x_{t-1}, a_{t-1}, e^*) Pr(a_t | x_{t-1}, a_{t-1}, e^*) = Pr(x_t, a_t | x_{t-1}, a_{t-1}, e^*) = Pr(w_t | w_{t-1}, e^*)
$$

Since a single equilibrium is played in a market-level time series, putting cross-sectional markets together results in the following mixture feature by the law of total probability and the first-order Markov property of $w_t$.

$$
Pr(\{w_{\tau}\}_t^T) = \sum_{e^*} Pr(\{w_{\tau}\}_t^T | e^*) Pr(e^*) = \sum_{e^*} \prod_{\tau=t}^{T-1} Pr(w_{\tau+1} | w_{\tau}, e^*) Pr(w_{\tau} | e^*) Pr(e^*)
$$

(13)

In this dynamic setup, the correlation between states in consecutive periods comes not only from the latent equilibrium but also directly from the Markov transition. Specifically, the correlation between $w_t$ and $w_{t+2}$ comes from two sources: $w_{t+1}$ and the underlying equilibrium. Note that the identification in the measurement error literature requires measurements are independent conditional on the latent variable. This condition can be satisfied if $w_t$ and $w_{t+2}$, when fixing $w_{t+1}$, are independent conditional on the underlying equilibrium. Consequently, three periods of data are needed for identification of the number of equilibria. With three periods of data, the joint distribution of $w_t$ becomes

$$
Pr(w_{t+2}, w_{t+1}, w_t) = \sum_{e^*} Pr(w_{t+2} | w_{t+1}, e^*) Pr(w_{t+1}, w_t | e^*) Pr(e^*)
$$

Fixing $w_{t+1}$, I rewrite the above equation into a matrix representation using variation of only $w_t$ and $w_{t+2}$.

$$
F_{w_{t+2}, w_{t+1}, w_t} = A_{w_{t+2}, w_{t+1}} e^* D e^* A_{w_{t+1}, w_t} e^*
$$

This equation linking knowns with unknowns enables me to identify the number of equilibria using rank inequality, which I stated in the following lemma.

**Lemma 4.** The rank of the observed matrix $F_{w_{t+2}, w_{t+1}, w_t}$ serves as the lower bound for the number of equilibria, i.e., $Q \geq \text{Rank}(F_{w_{t+2}, w_{t+1}, w_t})$. Furthermore, the number of equilibria is identified, in particular, $Q = \text{Rank}(F_{w_{t+2}, w_{t+1}, s_t})$ if the following conditions are satisfied

1. $H > Q$
2. both matrices $A_{w_{t+2}, w_{t+1}} e^*$ and $A_{w_{t+1}, w_t} e^*$ have full rank
3. all $Pr(e^*)$ are positive

**Proof** Proof is exactly the same as in Lemma 1.
There are several advantages in dynamic games compared to static frameworks in the identification of the number of equilibria. First of all, in a dynamic framework, the identification can use all variation including actions of all players and observed characteristics because the equilibrium is defined over all states. Thus, condition (1) holds easily. In contrast, static games can only rely on the variation from a part of the player’s actions because the equilibrium is characterized conditional on game characteristics. Secondly, the full rank condition means that enough variation in the conditional choice probability of different equilibria is needed to disentangle CCPs of each equilibrium. That is, not a single equilibrium is redundant. In the dynamic setting, the full rank condition is required for a fixing value of \( w_{t+1} \). Given the variation in the state space, it is easier to find a \( w_{t+1} \) so that the full rank condition is satisfied. Moreover, with more variation in the measurement, the full rank condition is easily satisfied too.

Note that the first order Markov property is guaranteed by the MPE assumption and only previous actions are included in the current state variable. If more periods of dependence are allowed either through the payoff function or state transition process, \( w_t \) will follow a higher-order of Markov process. Thus a longer period of data is needed for identification because of the longer lagged dependence.

Provided that the number of equilibria is identified, now I am going to identify all the CCPs of each equilibrium. With the inter-temporal dependence, using only three periods of data is not sufficient for the identification anymore. To control for the inter-temporal dependency, one extra period of data is required to control for the correlation between different periods. Specifically, if four periods of data are observed as \( w_t, w_{t+1}, w_{t+2}, w_{t+3} \), when \( w_{t+1} \) and \( w_{t+2} \) are fixed, the covariance between \( w_t \) and \( w_{t+3} \) comes from the underlying equilibrium. Otherwise they are independent. The joint distribution becomes:

\[
Pr(w_{t+3}, w_{t+2}, w_{t+1}, w_t) = \sum_{e^*} Pr(w_{t+3}|w_{t+2}, e^*) Pr(w_{t+2}|w_{t+1}, e^*) Pr(w_{t+1}, w_t|e^*) Pr(e^*) \quad (14)
\]

With the number of equilibria \( Q \) identified, I partition the state space \( \mathcal{W} \) from a dimension of \( H \) to a state space \( \tilde{\mathcal{W}} \equiv \{1, 2, ..., Q\} \) with a dimension of \( Q \) via a surjective function mapping \( G : \mathcal{W} \rightarrow \tilde{\mathcal{W}} \) and denote them as \( \tilde{w}_\tau = G(w_\tau) \) where \( \tau = t, t+3 \). Fixing \( w_{t+2} \) and \( w_{t+1} \), I rewrite equation 14 into the following matrix expression:

\[
F_{\tilde{w}_{t+3}, \tilde{w}_{t+2}, \tilde{w}_{t+1}, \tilde{w}_t} = A_{\tilde{w}_{t+3}|\tilde{w}_{t+2}, e^*} D_{\tilde{w}_{t+2}|\tilde{w}_{t+1}, e^*} D_{e^*} A_{\tilde{w}_{t+1}, \tilde{w}_t|e^*} \quad (15)
\]

Evaluating \( w_{t+2} \) to another value \( \tilde{w}_{t+2} \), I can obtain another matrix expression that share the common term \( A_{\tilde{w}_{t+1}, \tilde{w}_t|e^*} \) with equation 15. By the same logic, changing the value of \( \tilde{w}_{t+1} \), evaluating \( w_{t+2} \) in two different values results in two matrix equations sharing the common term \( A_{\tilde{w}_{t+1}, \tilde{w}_t|e^*} \). Using this feature,
manipulation over these four matrix expression leads to a matrix eigen-decomposition expression. As a result, the Markov law of motion \( Pr(w_{t+3}|w_{t+2}) \) can be identified using four periods of data, stated in the following lemma.

**Lemma 5. (Markov Law of Motion)** With four periods of data, distinctive eigenvalues and conditions in lemma 4 satisfied, the Markov law of motion \( Pr(w_{t+3}|w_{t+2}, e^*) \), the initial condition \( Pr(w_t|e^*) \), the strategies of each equilibrium \( Pr(a_t|x_t, a_{t-1}, e^*) \), the state transition \( Pr(x_t|x_{t-1}, a_{t-1}, e^*) \) and the equilibrium selection are uniquely identified through eigen-decomposition.

**Proof** See Appendix B

Now I am moving to identifying the expected period payoff using the equilibrium condition in dynamic setups (see Bajari, Chernozhukov, Hong, and Nekipelov (2009)). With identification of the expected period payoff defined as \( \Pi_i(a_i, s) = \sum_{a_{-i}} \pi_i(a_i, a_{-i}, s)\sigma_{-i}(a_{-i}|s) \), identification of payoff functions \( \pi_i(a_i, a_{-i}, s) \) is exactly the same as in the static case with exclusion restriction. Here since identification of the payoff functions can be done with any set of equilibria, for ease of notation, I drop the equilibrium notation. Since multiple equilibria indicate more constraints for the payoff functions, the presence of multiple equilibria helps for identification and estimation.

Note that the equilibrium condition with extreme value distribution assumption on the private shocks becomes:

\[
\log(\sigma_i(a_i = k, s)) - \log(\sigma_i(a_i = 0, s)) = V_i(a_i = k, s) - V_i(a_i = 0, s)
\]  

(16)

Define the ex ante value function for player \( i \) before obtaining private shocks, as

\[
V_i(s) = E_{\epsilon_i} \max_{a_i} V_i(a_i, s) + \epsilon_i(a_i) = \log\left(\sum_{k=0}^{K} \exp(V_i(k, s))\right)
\]

\[
= \log\left(\sum_{k=0}^{K} \exp(V_i(k, s) - V_i(0, s)) + V_i(0, s)\right)
\]

(17)

The second equation holds because of i.i.d. and the extreme value distribution assumption of \( \epsilon_i \). By definition of the choice specific value function \( V_i(a_i, s) \), we can relate the ex ante value function and
choice specific value function through the following equation:

\[ V_i(a_i, s) = \Pi_i(a_i, s) + \beta E(V_i(s')|s, a_i) \]

\[ = \Pi_i(a_i, s) + \beta E(\log \sum_{k=0}^{K} \exp(V_i(k, s) - V_i(0, s))) + V_i(0, s')|s, a_i) \]

\[ = \Pi_i(a_i, s) + \beta E(\log \sum_{k=0}^{K} \exp(V_i(k, s) - V_i(0, s))) + \beta E(V_i(0, s')|s, a_i) \]  \hspace{1cm} (18)

From the discrete choice model, as in static framework, it is impossible to identify all of the payoff primitives, so the same normalization condition in as the static case is imposed. That is, the payoff of choosing action 0 is normalized to be 0, i.e., \( \pi_i(a_i = 0, a_{-i}, s) = 0 \). For this special action \( a_i = 0 \), we have

\[ V_i(0, s) = \Pi_i(a_i = 0, s) + \beta E(\log \sum_{k=0}^{K} \exp(V_i(k, s) - V_i(0, s))) + \beta E(V_i(0, s')|s, a_i = 0) \]

\[ = \beta E(\log \sum_{k=0}^{K} \exp(V_i(k, s) - V_i(0, s))) + \beta E(V_i(0, s')|s, a_i = 0) \]  \hspace{1cm} (19)

By equilibrium condition 16, the term \( \log(\sigma_i(a_i = k, s)) - \log(\sigma_i(a_i = 0, s)) \) can be computed through \( \log(\sigma_i(a_i = k, s)) - \log(\sigma_i(a_i = 0, s)) \), so we can take it as a constant. Consequently, equation 19 provides a contract mapping on unknowns \( V_i(0, s) \). By Blackwell’s condition, a unique fixed point is guaranteed, so \( V_i(0, s) \) is identified. With identification of \( V_i(0, s) \), together with equation 17, the ex ante value function \( V_i(s) \) is identified. From equation 19, \( V_i(a_i = k, s) \) is identified. Then we can identify the expected period payoff \( \Pi_i(a_i, s) \) through equation 18 as:

\[ \Pi_i(a_i, s) = V_i(a_i, s) - \beta E(\log \sum_{k=0}^{K} \exp(V_i(k, s) - V_i(0, s))) - \beta E(V_i(0, s')|s, a_i) \]  \hspace{1cm} (20)

With the CCPs under each equilibrium identified, the identification of payoff functions is back to the traditional case with a unique equilibrium. The identification is provided in the Appendix. Identification proceeds first to identify the difference of choice specific current utility. Then with the exclusion restriction, the payoff functions \( \Pi_i(a_i, s) \) can be nonparametrically identified with a normalization assumption. Again, the existence of multiple equilibria can serve as an exclusion restriction, which relaxes the requirement of variation of the extra exclusion restriction.

**Theorem 3.** (Identification of Dynamic Games with Incomplete Information) With the conditions in lemma 4 satisfied, assumptions of MPE and distinctive eigenvalue, assumptions 3 and 4, the number of equilibria \( Q \), the equilibrium selection \( Pr(e^*) \), the strategies of each player in each equilibrium
Pr(a_i|a_{t-1}, x_t, e^*) and the payoff function π_i(a_i, a_{-i}, x) are nonparametrically identified in dynamic games with four periods of data.

6 Monte Carlo Simulation

This section presents some Monte Carlo evidence for the proposed identification and estimation methodology in the static game framework. Using the simulated data, first I estimate the number of equilibria, the CCPs of each equilibrium and the equilibrium selection mechanism. Then I estimate the payoff primitives.

Suppose n players decide to enter or stay out of markets with characteristics s ∈ S. Assume the market attribute is discrete. Thus, market characteristics s can be regarded as the market type. Assume players are homogeneous, and the payoff functions of entry (1) or not (0) are parameterized as follows:

\[ \pi(a_i = 1, a_{-i}; s) = \alpha s + \beta \sum_{a_{-i}} I(a_{-i} = 1) \frac{n-1}{n} + \epsilon_{i1} \]
\[ \pi(a_i = 0, a_{-i}; s) = \beta \sum_{a_{-i}} I(a_{-i} = 0) \frac{n-1}{n} + \epsilon_{i0} \]

where \( \epsilon_{i1} \) and \( \epsilon_{i0} \) are private shocks. Assuming the private information is independent and identically follows extreme value distribution. Given this specific payoff function, the number of players does not affect the equilibrium strategy, i.e., only the fraction of players entering matters. For \( \alpha = 0.04 \), \( \beta = 2.5 \) and only considering symmetric equilibria, all the equilibria for different market types \( s = 1, 2, 3, 4 \) are presented in figure 3. Regarding market types \( s = 1, 2, 3 \), there are three equilibria, among which the middle one is unstable, while there is a unique equilibrium for market type \( s = 4 \).

The Monte Carlo experiment consists of repetition of 500 with sample size of 1000 for each market type. For each replication, I generate the equilibrium according to the equilibrium selection mechanism, then generate the actions of each player according to the corresponding equilibrium strategies.

Estimation of the number of equilibria. The number of equilibria is estimated by a sequential test using the determinant of the associated matrix as the statistic. Note that the estimation of the rank is consistent as long as the significance level for the sequential testing approaches zero with a certain rate. I illustrate this asymptotic property through presenting the frequency of selecting the right number of equilibria with different numbers of sample size \( N = 200, 500, 1000, 2000 \) and \( 5000 \). The reported results are based on 1000 simulated samples from mixtures with two equilibria.
figure 2, the frequency of selecting the right number of equilibria approximates one as the sample size goes to infinity. Consequently, the estimation of the rank of a general square matrix is consistent.

**Estimation of CCPs of each equilibrium** To estimate the CCPs of each equilibrium, first I select the number of equilibria. Thus, estimation of the number of equilibria is a model selection procedure. The selection is consistent in large samples, however, the probability of choosing the wrong number of equilibria is positive in finite samples. With selecting the wrong number of equilibria, I obtain different sets of CCPs, making the comparison of CCPs to the true CCPs to be impossible. To investigate the performance of the estimation through eigenvalue-eigenvector decomposition, I assume that the number of equilibria is known. The estimation results in both cross-sectional and panel data structures show that the methodology manages to provide good estimates (see table 1 and table 3). CCPs are estimated with high accuracy, and so is the equilibrium selection probability.

**Structural parameters: unique versus multiple** To better understand the influence of the unique equilibrium assumption, I estimate the game primitives considering multiple equilibria and assuming a unique equilibrium. Estimation is through minimum distance based on the CCPs estimated above, or CCPs computed directly from the data when I assume uniqueness. Estimation results are presented in table 2.

The unique equilibrium assumption is problematic when multiplicity is presented. First of all, estimates are biased if one assumes only a unique equilibrium in the data. Moreover, I obtain an estimate with opposite sign to the true parameter, which makes the inference in the wrong direction. Things will become worse if one relies on the estimation result for policy regulation. For example, in this simple framework, firms are better off entering markets with bigger $s$, which is very intuitive because the bigger the market, the better. The unique equilibrium assumption, however, yields a confusing estimates with negative market effect.

The unique equilibrium assumption is not bad if we look at the estimates of the interaction effect. Still, the estimates are further from the true parameter compared to the estimates considering multiplicity, but the sign is correct and the bias is within a reasonable range. Again this is just one simple example, but it provides us some idea that avoiding the multiple equilibria issue will introduce estimation error. How big the error is depends on the whole environment, especially the equilibrium selection mechanism. For example, if players utilize one typical equilibrium most of the time, then assuming a unique equilibrium is not a bad approximation to the reality. However, without tackling this multiplicity issue, there is no outside information to judge whether this is the case. Consequently,
when doing empirical studies, we should be aware of the presence of multiple equilibria and be cautious of making the unique equilibrium assumption.

**Testing: multiple equilibria versus payoff-relevant latent states** As I discuss above, both multiple equilibria and payoff-relevant latent states yield mixture features in the actions data. For comparison, assume the support of the latent variable is finite and fixed across different markets. Under the null hypothesis that only multiple equilibria are presented, estimation of the game primitives $\theta$ using CCPs of any equilibrium set should yield the same estimates.

Suppose the econometrician observes $s$ and the actions players choose, only multiple equilibria are presented in the data. To do the test, I estimate two sets of structural parameters through one set of CCPs ($\{Pr(a_i = 1|s = 1, 2, 3, \tau = 1), Pr(a_i = 1|s = 1, \tau = 2), Pr(a_i = 1|s = 4)\}$) and another set of CCPs($\{Pr(a_i = 1|s = 1, 2, 3, \tau = 2), Pr(a_i = 1|s = 2, \tau = 1), Pr(a_i = 1|s = 4)\}$) separately. Using CCPs from both equilibria of some values of $s$ is for identification purposes. When players are identical, the parameters are not identified without multiplicity. The testing fails to reject the null hypothesis that multiple equilibria are presented in the data. No other payoff-relevant latent variable exists.

Assume that the market characteristics are unobserved by the econometrician. To make the overall dimension of unobserved factor not too big, here I only consider markets $s = 1, 4$, among which there are two equilibria selected in market $s = 1$. Consequently, the dimension of the unobserved element is combined to be 3, which is assumed to be known. With matrices decomposition, three sets of CCPs are estimated. Using any two combination of the CCPs, one can estimate the structural parameters $\beta$, and test whether this $\beta$ is the same or not. The testing rejects the null hypothesis that the structural primitives estimated are the same.

### 7 Empirical Application: Commercial Break Decisions by Stations

This section applies the proposed estimation methodology to commercial timing decisions by stations with contemporary music formats (Contemporary Hit Radio (CHR)/Top 40, Country, Rock etc.), and provides empirical evidence of multiple equilibria. Specifically, there are two equilibria in which stations cluster to one time slot to air their commercials.
7.1 Institution Background and Data

Listeners dislike listening to commercials so they seek to avoid listening by switching to other stations or outside options such as tapes or CDs while commercials are on the air. Advertisers for sure prefer stations to play their commercials at the same time to reduce commercial avoidance. Stations, on the other hand, may have different incentives because values of commercials are not based on listenership of a particular commercial. In reality, average commercial audiences are not measured. Instead, Arbitron, the radio ratings company, estimates a station’s average audience by averaging over both commercial and noncommercial programming. As a result, the average audience might increase if stations play commercials at different times to keep listeners tuned in to the radios stations instead of seeking outside options.

Stations tend to play commercials at the same time. Specifically, figure 3 presents the average proportion of stations playing commercials in each minute during two different hours of the day, and those commercial timing distributions are far from uniform. One possible explanation is that coordination increases station’s commercial values. Another possible reason, however, is the existence of common factors which makes different time slots of each hour particularly desirable for commercials. Knowledge of the station industry indicates that common factors do affect timing decisions. For example, the way that Arbitron computes listenership strongly affects station’s commercial break decisions.

Common factors, however, are not a perfect explanation for the clustering behavior. Suppose common factors are indeed the underlying drive for the clustering. Note that Arbitron uses the same methodology to compute the listenerships. Consequently, one can expect that commercials are clustered in every market, and also at the same times across markets. This is not the case in reality. See figure 4 for stations in two markets playing commercials during one particular hour. The distributions of commercial breaks in both markets have three peaks, which are at noticeably different times. A similar situation exists in the aggregate distribution. The clustering of commercials at different times in different markets is not driven by unobserved common factors.

Another possible explanation is the presence of multiple equilibria. In the static game, stations strategically choose times to air their commercials. Stations coordinate to air their commercials at the same time to avoid listener switching. Multiple equilibria are presented and different equilibria are employed across markets. This rationalizes both the coordination in one market and the different times of the clustering across markets. With this static setup with the possibility of multiple equilibria, this
paper uses the methodology presented above to investigate how many equilibria are actually in the data.

The data used in this paper is the same dataset as that in Sweeting (2006), which constructs the data on the timing of commercials by music radio stations in US metro markets using hourly airplay logs collected by Medabase 24/7. The data is extracted from airplay logs that record the music that stations play on a minute-by-minute basis. In summary, there are 144 markets in total. The number of stations in each market varies from 2 to 20 with a mean of 13. Stations not choosing either action are excluded from the estimation. Each station has 236 observations, including 59 days, and each day we have 4 different hour timing choices with two being drive-time and two non drive-time. From the summary statistics 4, the proportion of stations choosing option 1 is slightly greater than that choosing option 0. For a detailed description of the data, refer to ?.

7.2 Model Setup

While actual commercial timing is continuous, a discrete feature exists in the schedule of commercials on music stations, because timing decisions involve planning the order of songs and commercial breaks. For example, the programmer considers the commercial breaks in the gaps between the songs. As a result, stations are modeled to choose playing their commercials in finite time blocks simultaneously. Stations can play several sets of commercials at many different times during an hour. Estimation of games with this features is beyond the current literature and the scope of this paper. Instead, the choice of commercials breaks by stations is modeled as a simple binary choice game. Specifically, I use information about whether commercials are being played at two particular times each hour, :48–:52 and :53–:57, denoted as option 0 and option 1 respectively.

Assume that stations are identical, so individual stations do not have station characteristics. Following Sweeting (2006), station $i$’s payoff for placing a commercial in time block $t \in \{0, 1\}$ is defined as follows:

$$\pi_{it} = \alpha_t + \beta P_{-it} + \epsilon_{it}$$

where $p_{-it}$ is the proportion of stations in the same market choosing timing $t$. $\alpha_t$ allows different average profit for stations when they play their commercials in different timing $t$, and $\beta$ captures the coordination incentives. Stations receive the idiosyncratic private profit shocks before they make their timing decisions. The $\epsilon$s represent the fact that a station tends to play commercials at different times every day are represented by $\epsilon_{it}$. This variation is because the length of other programming, such as
As usual, I assume $\epsilon_{it}$ to be i.i.d with extreme value distributions across actions, players and markets. $\alpha_0$ is normalized to be zero for identification purposes. Denote $\alpha_1$ as $\alpha$ for ease of notation. With the payoff specification, the number of stations within each individual market does not matter. Thus, information from different markets can be pooled for estimation. Note that without multiple equilibria, the model is under-identified because there is one equation with two unknowns from the equilibrium condition. Exclusion restrictions do not apply here because stations are identical. If there are at least two equilibria, the proportion of players other than player $i$ choosing action 1 is different under different equilibria. Thus, the coordination effect $\beta$ is identified, and $\alpha$ is identified.

### 7.3 Empirical Results

This section presents the estimation results. Estimation results show that two equilibria exist and stations stick to the same equilibrium over time. These results are consistent with that of Sweeting (2006).

Even though I have panel data, I do not make any assumptions about the equilibrium employed over time by the same market. For instance, markets employ the same equilibrium over time. Treating markets in different days as different markets, panel data can be constructed to be a cross-sectional. Secondly, to investigate whether markets stick to the same equilibrium or not over time, I assume that the equilibrium employed over time follows a first-order Markov chain. Note that this assumption nests the case of the same equilibrium employed over time by an identity transition matrix. The approach used here is different from Sweeting (2006), who assumes that the same equilibrium is employed over time by the same market. Thus, this empirical attempt provides us with some idea about whether making the same equilibrium assumption has empirical support or not.

**Big Versus Small Markets** The number of stations varies in different markets. The bigger the market, the more stations it has. The more stations, the harder it becomes to coordinate. To control the market effect but still pool data from different markets together, I divide markets into two types, big versus small, according to the rank by population.

Not surprisingly, there is a unique equilibrium in big markets due to more difficult coordination. In contrast, I find two equilibria in small markets. The two equilibria are similar to each other in that
the probability of clustering in each time interval is relatively similar in each equilibrium. Interestingly, players are inclined to employ one equilibrium more often than the other, with a probability of 0.73 versus 0.27.

The $\alpha$ coefficient measures whether time block :53-:57 is more attractive for commercials independent of any incentive to coordinate. I fail to reject the null hypothesis that $\alpha$ is significantly different from zero. This insignificance is consistent with the fact that both time intervals are equally distant from the quarter-hours, which are known to be unattractive times for commercials.

I get similar estimates when I utilize the panel data structure. Two equilibria exist in small markets. Moreover, the estimated CCPs and the structural parameters are similar for both cross-sectional and panel data. Again, one cannot reject the null hypothesis that the $\alpha$s equals zeros, indicating that neither of the options is more attractive than the other. Moreover, markets stick to the same equilibrium over time because the probability of employing the same equilibrium in the previous day does not significantly differ from 1. This interesting finding suggests that time series data might be free of multiple equilibria.

**Drive-time Versus Non Drive-time** Commercials are planned in every hour, and different hours are treated differently. Thus I treat data from different hours as separate markets. There are big differences between drive-time and non drive-time. In-car listeners are more likely to switch stations to avoid commercials than those at home or at work, and there are more proportional in-car listeners during drive-time. How strong the incentive is to coordinate depends on how much listeners dislike commercials and how easy it is for them to switch stations. For example, listeners might respond to commercials differently during driving time and non-driving time, because during driving time they stay in cars and it is very easy to switch stations. For example, in-car listeners switch stations 29 times per hour on average to avoid commercials (McDowell and Dick (2003)).

As expected, the strategic interaction is stronger during driving-time, and multiple equilibria only exist during driving time. The presence of multiple equilibria during drive-time but not outside drive-time is consistent with the model. Strong incentive to coordinate is the reason to have multiple equilibria, and the incentive should be greater during drive-time when there are more in-car listeners.

I index equilibrium 1 as the equilibrium that stations cluster at option 0 (:48-:52), meaning the probability of choosing option 0 is greater than one half. In all estimation with different sub-samples, one can see that the probability of employing equilibrium 1 is less than half, suggesting that stations slightly prefer equilibrium 2, which clusters on option 1(:53-:57). Moreover, the equilibrium selections in
both drive-time hours are similar, with a probability of choosing equilibrium 1 to be 0.3307 versus 0.4192 for 4-5 pm and 5-6 pm respectively. This result at least provides us with supports that sometimes it is not a bad assumption that the equilibrium selection mechanisms are the same across different markets.

8 Conclusion

I have developed a methodology to nonparametrically identify finite games with incomplete information allowing the presence of multiple equilibria. In particular, I show that the number of equilibria, strategies of players in each equilibrium and the equilibrium selection mechanism are identified in both static games and dynamic games. Payoff primitives can be identified using exclusion restrictions. A Monte Carlo evidence shows that the estimators perform well in median-size samples. As an application of the proposed methods, I study the behavior of stations which strategically choose their time break to air commercials using the same dataset used in Sweeting (2006). I find out that two equilibria are employed in smaller markets with stations clustering in one time slot to air commercials in one equilibrium. Moreover, about half of the markets employ one equilibrium. In addition, markets stick to the same equilibrium over time.

The existence of multiple equilibria and common unobserved heterogeneity are observable equivalent in terms of the mixture feature. However, assuming common unobserved heterogeneity cannot rule out the presence of multiple equilibria. Instead, it makes the identification more difficult because now two types of latent variables mixed together. Identification of dynamic games with both multiple equilibria and common unobserved heterogeneity is similar since multiple equilibria and unobserved heterogeneity are the same across all states. Moreover, distinguishing the two latent variable relies on the fact the multiple equilibria map to the same payoff functions but unobserved heterogeneity map to different payoff functions. In static games, identification is more difficult because the BNE is defined conditional on the observed state variables. This means that it is possible the dimension of the multiple equilibria varies across different observed state variables. On the other hand, the dimension of common unobserved heterogeneity is usually assumed to be fixed across different observables. Consequently, identification of the payoff functions needs to distinguish the CCPs from different common unobserved heterogeneity, which is not clear right away. As a result, I study the identification of both static and dynamic games with incomplete information allowing both multiple equilibria and common unobserved heterogeneity in a subsequent paper.
The BNE is defined conditional on the state variable in static games while MPE considers all possible states together because of the dynamic feature. As a result, the number of equilibria varies when markets have different state variables in static games, contrasting to the fact that the number of equilibria is invariant to the values of the state variables in dynamic games. This is because players are forward looking in dynamic games. To make better decision, they consider all possible situations.

References


Appendix

The appendix includes four sections, tables and figures. Section A provides the proof of lemmas and propositions. Section B presents the identification of payoff functions in both static and dynamic games.

A Why Incomplete Information

First of all, games with incomplete information results in a global mixture feature while games with complete information have a local mixture. Specifically, in a game with incomplete information, the joint distributions of players satisfies the following equations:

\[
Pr(a_1, a_2) = \sum_{e^*} Pr(a_1, a_2|e^*)Pr(e^*) \\
= \sum_{e^*} Pr(a_1|e^*)Pr(a_2|e^*)Pr(e^*)
\]

While in a complete information game, we have

\[
Pr(a_1, a_2) = \int_{u_1, u_2} Pr(a_1, a_2|u_1, u_2)f(u_1, u_2)du_1du_2 \\
= \int_{u_1, u_2} \sum_{e^*} Pr(a_1, a_2|u_1, u_2, e^*)Pr(e^*|u_1, u_2)f(u_1, u_2)du_1du_2 \\
= \int_{u_1, u_2} \sum_{e^*} Pr(a_1|u_1, u_2, e^*)Pr(a_2|u_1, u_2, e^*)Pr(e^*|u_1, u_2)f(u_1, u_2)du_1du_2
\]

(A.1)

To use the identification technique in measurement error literature, the observed actions of different players need to be independent conditional on the latent variable. One can see that the conditional independence condition is satisfied when conditional on \(u_1\) and \(u_2\), but the independence property disappears when aggregate over all the profit shocks. One idea to regain this nice conditional independence is that instead of integrating over \(u_1\) and \(u_2\), we can perceive \(u_1\) and \(u_2\), together with the equilibrium, as a new latent variable. To see this idea clearly, suppose \(u_1\) and \(u_2\) are discrete, then the above equation A.1 becomes:

\[
Pr(a_1, a_2) = \sum_{u_1, u_2, e^*} Pr(a_1|u_1, u_2, e^*)Pr(a_2|u_1, u_2, e^*)Pr(u_1, u_2, e^*) \\
= \sum_{\tau} Pr(a_1|\tau)Pr(a_2|\tau)Pr(\tau)
\]
where $\tau$ is a variable that collects all the possible values of $u_1$ and $u_2$ and the equilibrium $e^*$. Note that the dimension of $e^*$ varies with $u_1$ and $u_2$. This essentially is the same case as there are both unobserved payoff-relevant heterogeneity and multiple equilibria in incomplete information games. Since during the eigenvalue-eigenvector decomposition, it is hard to rank the eigenvalues using the value of $Pr(a_1|\tau)$. The reason is that we cannot tell whether different values of $\tau$ come from different values of $u_1$ and/or $u_2$, or it comes from the same $u_1$ and $u_2$ but different equilibria. Specifically, $u_1$ and $u_2$ are both binary with $\{0, 1\}$, hence we have the combination of $u_1$ and $u_2$ denoted as $u=\{1=[0 0], 2=[0 1], 3=[1 0], 4=[1 1]\}$. Assume that there is a unique equilibrium for $u = 1, 2$, but there are two equilibria for $u = 3, 4$. Thus, $\tau$ is a dimension of 6, indexed as $1, 2, 3, 4, 5, 6$. Suppose there is enough variation that we can identify the support of the latent variable to be 6. With that, from eigenvalue-eigenvector decomposition we can identify six set of $Pr(a_1|\tau)$. Now the problem is that which two $\tau$ come from the same value of $u$.

Moreover, unlike an incomplete information game, a complete information game does not have a nice closed form equilibrium equation. The equilibrium condition is through a comparison of expected payoffs from different actions. Different set of equilibria satisfy different inequalities. Thus, if I manage to distinguish those eigenvalues that which one comes from which $u$ and $e^*$, it is still hard to identify the structural parameters. In the case of unique equilibrium, there is no problem of mapping the equilibrium predicted by theory and the one identified from the decomposition. In the case of multiple equilibria, another issue arises is how can we tell from the theory which equilibrium is the one that we identify from the data. Moreover, it is not necessary that players employ all the equilibria in reality. In the incomplete information game, we do not need to worry about this issue because all the equilibria satisfy the same equilibrium condition. To identify the payoff functions, we just need to plug the identified equilibrium back into the same equilibrium equations.

To sum up, first of all, complete information games provide us a local mixture feature while incomplete information games provides a global mixture. Secondly, with treating the private shocks and the latent equilibrium as a new latent variable, we have a global mixture but with two different kinds of latent variable. Consequently, problems exist as to ranking the eigenvalues during decomposition. Finally, even if we are free of ordering issues, without a nice closed form equilibrium condition, a problem arises regarding how to map the equilibrium identified from the data to the ones predicted by theory when there are multiple equilibria on the way to identifying the payoff structural parameters.
B Proofs

Proof of Lemma 1  Based on conditional independent assumptions, the joint distribution of actions from two players can be expressed as:

\[ Pr(s_1, s_2) = \sum_{e^*} \sigma_1(s_1|e^*)\sigma_2(s_2|e^*)Pr(e^*) \]

Rewrite it into a matrix form:

\[ F_{s_1,s_2} = A_{s_1|e^*}DA_{s_2|e^*}^T \]

With assumptions that \((K + 1)^l > Q\) and full rank of both matrices \(A_{s_1|e^*}\) and \(A_{s_2|e^*}\), then according to the following inequality regarding the rank of matrix \(F_{s_1,s_2}\):

\[
\text{Rank}(A_{s_1|e^*}) + \text{Rank}(A_{s_2|e^*}) - Q \leq \text{Rank}(F_{s_1,s_2}) \leq \min\{\text{Rank}(A_{s_1|e^*}), \text{Rank}(A_{s_2|e^*})\} \quad (B.1)
\]

I conclude that \(\text{Rank}(F_{s_1,s_2}) = Q\). ■

Proof of Lemma 2  Matrix A is with dimension of \(K + 1 \times K + 1\), and rank of matrix A equals \(Q\). Without loss of generality, assume that \(K = Q\) and denoted \(A\) as \([a_1, ..., a_Q, a_{Q+1}]\) where \(a_i\) are row vectors. Since \(\text{rank}(A) = Q\), among the \(Q + 1\) column vectors, there at least exists \(Q\) of them that are linearly independent. Again, w.l.o.g, assume \(a_1, ..., a_Q\) are linearly independent. By the definition of linear independence, there exists a series of \(\lambda_1, ..., \lambda_Q\) such that

\[ a_{Q+1} = \lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_Q a_Q \]

where there must exist a \(\lambda_i \neq 0\). Moreover, \(\lambda_i > 0\) because \(a_1, ..., a_K\) are all positive by the nature of probability. Next I prove that partitioning row vector \(i\) and \(Q + 1\) to be a new group results in a linear independent \(Q\) vectors. That is, the new \(Q\) row vectors \(a_1, a_2, ..., a_i + a_{Q+1}, a_{i+1}, ..., a_Q\) are linearly independent. To prove the linear independence, we need to prove that for any \(\eta_1, ..., \eta_Q\) satisfying

\[ \eta_1 a_1 + \eta_2 a_2 + ... + \eta_i (a_i + a_{Q+1}) + ... + \eta_Q a_Q = 0 \]

we have \(\eta_1 = \eta_2 = ... = \eta_Q = 0\). Plug \(a_{Q+1} = \lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_Q a_Q\) back into above equation, leading to:

\[ (\eta_1 + \eta_i \lambda_1) a_1 + (\eta_2 + \eta_i \lambda_2) a_2 + ... + \eta_i (1 + \lambda_i) a_i + ... + (\eta_Q + \eta_i \lambda_Q) a_Q = 0 \]

Given that \(a_1, ..., a_Q\) are linearly independent, all the linear coefficients of the above linear combination should equal zero. Thus, \(\eta_k + \eta_i \lambda_k = 0, \forall k = 1, ..., Q\) and \(\eta_i (1 + \lambda_i) = 0\). Given that \(\lambda_i > 0\) by assumption,
η_i(1 + λ_i) = 0 implies that η_i = 0. Then η_k = 0 for k = 1,...Q. Thus, a_1, a_2, ..., a_i + a_{Q+1}, a_{i+1}, ..., a_Q are linear independent.

**Proof of Lemma 3**  With eigenvalue-eigenvector decomposition, matrix \( D_{a_3=k|e^*} \) and \( A_{\tilde{a}_1|e^*} \) are identified. Varying actions \( a_3 \) for player 3, the main equation holds with the same matrix \( A_{\tilde{a}_1|e^*} \). Thus, I do not have to go through eigenvalue-eigenvector decomposition to obtain other \( D_{a_3|e^*} \) when \( a_3 \neq k \). For actions other than \( k \) that player 3 might select, \( D_{a_3|e^*} \) can be identified through

\[
D_{a_3|e^*} = A_{\tilde{a}_1|e^*}^{-1} F_{\tilde{a}_1,\tilde{a}_2,a_3=k}^{-1} A_{\tilde{a}_1|e^*}
\]

Now I continue to identify the equilibrium selection mechanism, which essentially is the diagonal elements in diagonal matrix \( D \). Similarly, I have \( Pr(\tilde{a}_1) = \sum Pr(\tilde{a}_1|e^*)Pr(e^*) \) with matrix representation \( F_{\tilde{a}_1} = A_{\tilde{a}_1|e^*} D_{e^*} \). Thus, the equilibrium selection mechanism can be identified through \( D_{e^*} = A_{\tilde{a}_1|e^*}^{-1} F_{\tilde{a}_1} \).

For player 1, \( A_{\tilde{a}_1|e^*} \) is identified as eigenvectors of the decomposition, but we are interested in \( A_{a_1|e^*} \). From the joint distribution of \( a_1 \) and \( g_2 \), we have:

\[
F_{\tilde{a}_1,\tilde{a}_2} = A_{a_1|e^*} D_{e^*} A_{a_2|e^*}^T
\]

Since \( D_{e^*} \) and \( A_{a_2|e^*}^T \) are identified and invertible, then \( A_{a_1|e^*} \) is identified. \( A_{a_2|e^*} \) can be identified in the same procedure.

**Proof of Theorem 2**  With the first-order Markov process assumption on the equilibrium evolution and total probability, the joint distribution of three periods of data can be represented as:

\[
Pr(a_{t+2}, a_{t+1}, a_t) = \sum_{e_t^*,e_{t+1}^*,e_{t+2}^*} Pr(a_{t+2}, e_{t+2}^*, a_{t+1}, e_{t+1}^*, a_t, e_t^*)
\]

\[
= \sum_{e_t^*,e_{t+1}^*,e_{t+2}^*} Pr(a_{t+2}|e_{t+2}^*) Pr(e_{t+2}^*|e_{t+1}^*) Pr(a_{t+1}|e_{t+1}^*) Pr(e_{t+1}^*|e_t^*) Pr(a_t|e_t^*)
\]

\[
= \sum_{e_t^*,e_{t+1}^*} Pr(a_{t+2}|e_{t+1}^*) Pr(a_{t+1}|e_{t+1}^*) Pr(a_t|e_{t+1}^*) Pr(e_{t+1}^*|e_t^*)
\]

\[
= \sum_{e_t^*,e_{t+1}^*} Pr(a_{t+2}|e_{t+1}^*) Pr(a_{t+1}|e_{t+1}^*) Pr(a_t|e_{t+1}^*) Pr(e_{t+1}^*)
\]

where \( prob(a_t|e_{t+1}^*) \) represents the probability of the players choosing action \( a_t \) in period \( l \) when the equilibrium chosen in period \( t \) is \( e_{t+1}^* \); \( prob(e_{t+1}^*) \) is the fraction of markets that employ equilibrium \( e_{t+1}^* \) at period \( t+1 \), i.e., the equilibrium selection mechanism. The above equation holds because the
only dynamic across time is through the transition of equilibria, so here we do not require the private information to be independent across different players.

Summing over \(a_{t+1}\) yields

\[
Pr(a_{t+2}, a_t) = \sum_{e^*_{t+1}} Pr(a_{t+2}|e^*_{t+1})Pr(a_t|e^*_{t+1})Pr(e^*_{t+1})
\]

Given that the number of equilibria is identified, partition the \((K+1)^n\) alternatives into a \(Q\) alternative according to lemma 3, and denote as \(\tilde{a}_t = \{b_{\tau_1}, ..., b_{\tau_Q}\}\) for \(\tau = t, t + 2\) so that the matrix \(F_{\tilde{a}_{t+2}, \tilde{a}_t}\) defined accordingly is invertible. Matrix representation for the joint probability distribution equations with the following matrices definitions, we have

\[
F_{\tilde{a}_{t+2}, \tilde{a}_t} = A_{\tilde{a}_{t+2} | e^*_{t+1}}DA^T_{\tilde{a}_t | e^*_{t+1}} \tag{B.2}
\]

\[
F_{\tilde{a}_{t+2}, \tilde{a}_t, a_{t+1}=k} = A_{\tilde{a}_{t+2} | e^*_{t+1}} D_{a_{t+1}=k} A_{\tilde{a}_t | e^*_{t+1}} \tag{B.3}
\]

Since matrix \(F_{\tilde{a}_{t+2}, \tilde{a}_t}\) is invertible, I can post-multiply \(A^{-1}_{\tilde{a}_{t+2}, \tilde{a}_t}\) into both sides of equation B.3, leading to the following main equation.

\[
F_{\tilde{a}_{t+2}, \tilde{a}_t, a_{t+1}=k}F^{-1}_{\tilde{a}_{t+2}, \tilde{a}_t} = A_{\tilde{a}_{t+2} | e^*_{t+1}} D_{a_{t+1}=k} A_{\tilde{a}_t | e^*_{t+1}} \tag{B.4}
\]

With the distinctive eigenvalues assumption stated below, the \(Pr(a_{t+1} = k|e^*)\) is identified as eigenvalues of matrix \(F_{\tilde{a}_{t+2}, \tilde{a}_t, a_{t+1}=k}F^{-1}_{\tilde{a}_{t+2}, \tilde{a}_t}\).

**Assumption B.1.** (Distinctive Eigenvalues) there exists one choice \(k\) in period \(t + 1\), i.e., \(a_{t+1} = k\) that for any two equilibria, \(i \neq j\), \(Pr(a_{t+1} = i|e^*_{t+1} = q) \neq Pr(a_{t+1} = i|e^* = k)\) the probability of this action taken under different equilibria is different.

Like the proof in proposition 1, all the CCPs in different equilibria can be identified. Since \(Pr(a_{t+1} | e^*_{t+1})\) is a joint distribution for the \(n\) players, CCPs for individual players \(Pr(a_{t+1}|e^*_{t+1})\) can be identified by summing over \(a_{t+1}\) on \(Pr(a_{t+1}|e^*_{t+1})\).

In addition, the equilibrium evolution probability satisfies the following equation:

\[
Pr(\tilde{a}_{t+2}|e^*_{t+1}) = \sum_{e^*_{t+2}} Pr(\tilde{a}_{t+2}, e^*_{t+2}|e^*_{t+1}) = \sum_{e^*_{t+2}} Pr(\tilde{a}_{t+2}|e^*_{t+2})Pr(e^*_{t+2}|e^*_{t+1})
\]

where \(Pr(e^*_{t+2}|e^*_{t+1})\) represents the probability of equilibrium \(e^*_{t+2}\) chosen in period \(t + 2\) when the equilibrium in period \(t + 1\) is \(e^*_{t+1}\), i.e., the equilibria evolution. Also \(Pr(\tilde{a}_{t+2}|e^*_{t+2})\) is identified as
eigenvectors of the decomposition with normalization by column sum equals to 1, and $Pr(\alpha_{t+2}e^*_t)$ is the same as $Pr(\alpha_{t+1}e^*_{t+1})$. Rewrite them into matrix form so that the equilibrium evolution process is identified.

**Proof of Lemma 5** With the surjective function $G$, fixing $w_{t+2}$ and $w_{t+1}$, matrix $F_{\bar{w}_{t+3},w_{t+2},w_{t+1},\bar{w}_t}$ is invertible. As a result, matrices $A_{\bar{w}_{t+3}|w_{t+2},e^*}$ and $A_{w_{t+1}|\bar{w}_t,e^*}$ are also invertible. Evaluating the joint distribution of four periods of data at four pairs of points $(w_{t+2}, w_{t+1}), (\bar{w}_{t+2}, w_{t+1}), (w_{t+2}, \bar{w}_{t+1}), (\bar{w}_{t+2}, \bar{w}_{t+1})$, each pair of equations will share one matrix in common. Specifically,

\begin{align}
  &\text{for } (w_{t+2}, w_{t+1}) : \
  &\quad \sum_{\bar{w}_t} F_{\bar{w}_{t+3},w_{t+2},w_{t+1},\bar{w}_t} = A_{\bar{w}_{t+3}|w_{t+2},e^*}D_{w_{t+2}|w_{t+1},e^*}e^*A_{w_{t+1}|\bar{w}_t,e^*} \\
  &\quad (B.5) \\
  &\text{for } (\bar{w}_{t+2}, w_{t+1}) : \
  &\quad \sum_{\bar{w}_t} F_{\bar{w}_{t+3},\bar{w}_{t+2},w_{t+1},\bar{w}_t} = A_{\bar{w}_{t+3}|\bar{w}_{t+2},e^*}D_{\bar{w}_{t+2}|w_{t+1},e^*}e^*A_{w_{t+1}|\bar{w}_t,e^*} \\
  &\quad (B.6) \\
  &\text{for } (w_{t+2}, \bar{w}_{t+1}) : \
  &\quad \sum_{\bar{w}_t} F_{\bar{w}_{t+3},w_{t+2},\bar{w}_{t+1},\bar{w}_t} = A_{\bar{w}_{t+3}|w_{t+2},e^*}D_{w_{t+2}|\bar{w}_{t+1},e^*}e^*A_{\bar{w}_{t+1}|\bar{w}_t,e^*} \\
  &\quad (B.7) \\
  &\text{for } (\bar{w}_{t+2}, \bar{w}_{t+1}) : \
  &\quad \sum_{\bar{w}_t} F_{\bar{w}_{t+3},\bar{w}_{t+2},\bar{w}_{t+1},\bar{w}_t} = A_{\bar{w}_{t+3}|\bar{w}_{t+2},e^*}D_{\bar{w}_{t+2}|\bar{w}_{t+1},e^*}e^*A_{\bar{w}_{t+1}|\bar{w}_t,e^*} \\
  &\quad (B.8)
\end{align}

Matrices $A_{\bar{w}_{t+3}|w_{t+2},e^*}$ and $A_{w_{t+1}|\bar{w}_t,e^*}$ are invertible by construction. Assume that $Pr(w_{t+2}|w_{t+1},e^*)$ is positive for every combination of $w_{t+2}$ and $w_{t+1}$, so matrix $D_{w_{t+2}|w_{t+1},e^*}$ is also invertible. Consequently, we can postmultiply inverse of equation B.6 to equation B.5, to obtain:

\begin{align}
  Y &= F_{\bar{w}_{t+3},w_{t+2},w_{t+1},\bar{w}_t}F_{\bar{w}_{t+3},\bar{w}_{t+2},w_{t+1},\bar{w}_t}^{-1} = A_{\bar{w}_{t+3}|w_{t+2},e^*}D_{w_{t+2}|w_{t+1},e^*}e^*D_{\bar{w}_{t+2}|w_{t+1},e^*}e^*A_{\bar{w}_{t+3}|w_{t+2},e^*} \\
  &\quad (B.9)
\end{align}

Similarly,

\begin{align}
  Z &= F_{\bar{w}_{t+3},\bar{w}_{t+2},w_{t+1},\bar{w}_t}F_{\bar{w}_{t+3},\bar{w}_{t+2},\bar{w}_{t+1},\bar{w}_t}^{-1} = A_{\bar{w}_{t+3}|\bar{w}_{t+2},e^*}D_{\bar{w}_{t+2}|\bar{w}_{t+1},e^*}e^*D_{\bar{w}_{t+2}|\bar{w}_{t+1},e^*}e^*A_{\bar{w}_{t+3}|\bar{w}_{t+2},e^*} \\
  &\quad (B.10)
\end{align}

Consequently, I postmultiply Eq. B.9 by Eq. B.10, leading to

\begin{align}
  YZ &= A_{\bar{w}_{t+3}|w_{t+2},e^*}[D_{w_{t+2}|w_{t+1},e^*}D_{\bar{w}_{t+2}|\bar{w}_{t+1},e^*}]^{-1}A_{\bar{w}_{t+3}|w_{t+2},e^*} \\
  &\equiv A_{\bar{w}_{t+3}|w_{t+2},e^*}D_{w_{t+2}|w_{t+1},e^*}D_{\bar{w}_{t+2}|\bar{w}_{t+1},e^*}D_{\bar{w}_{t+2}|w_{t+1},e^*}e^*A_{\bar{w}_{t+3}|w_{t+2},e^*} \\
  &\quad \text{where} \\
  &\quad D_{w_{t+2},\bar{w}_{t+2},w_{t+1},\bar{w}_{t+1}|e^*} = D_{w_{t+2}|w_{t+1},e^*}D_{\bar{w}_{t+2}|\bar{w}_{t+1},e^*}D_{\bar{w}_{t+2}|w_{t+1},e^*}D_{\bar{w}_{t+2}|w_{t+1},e^*} \\
  &\quad = Pr(w_{t+2}|w_{t+1},e^*)Pr(\bar{w}_{t+2}|\bar{w}_{t+1},e^*)Pr(w_{t+2}|w_{t+1},e^*)Pr(\bar{w}_{t+2}|\bar{w}_{t+1},e^*) \\
  &\quad \equiv C(w_{t+2}, w_{t+2}, w_{t+1}, w_{t+1}|e^*)
\end{align}

Equation B.11 results in a eigenvalue-eigenvector decomposition for observed matrix $YZ$, with eigenvectors corresponding to the matrix $A_{\bar{w}_{t+3}|w_{t+2},e^*}$ and the eigenvalues corresponding to matrix $D_{w_{t+2},\bar{w}_{t+2},w_{t+1},\bar{w}_{t+1}|e^*}$.

Again the eigenvalues need to be distinct as stated in the following assumption:
Assumption B.2. (Distinctive Eigenvalues) There exist \( w_{t+2}, \tilde{w}_{t+2}, w_{t+1}, \tilde{w}_{t+1} \), such that 
\[ C(w_{t+2}, \tilde{w}_{t+2}, w_{t+1}, \tilde{w}_{t+1}|e^* = i) \neq C(w_{t+2}, \tilde{w}_{t+2}, w_{t+1}, \tilde{w}_{t+1}|e^* = j) \] for any two equilibria \( i \neq j \).

This distinctive eigenvalues assumption is empirically testable because the matrix for the eigen-decomposition can be computed from the data. Moreover, the eigenvectors \( A_{\tilde{w}_{t+3}|w_{t+2},e^*} \) in the decomposition is unique up to multiplication by a scalar constant, which can be pinned down because each column should be summed up to one. Since the index of equilibria does not convey any economic meaning, so no additional monotonicity condition is needed for ordering of the eigenvalues. Therefore, for any given \( w_{t+2} \), \( Pr(\tilde{w}_{t+3}|w_{t+2}, e^*) \) is nonparametrically identified. Then \( A_{w_{t+2},w_{t+1}|e^*} \) can be identified using the joint distribution of \( \tilde{w}_{t+3}, w_{t+2}, w_{t+1} \) as follows:
\[
F_{\tilde{w}_{t+3},w_{t+2},w_{t+1}} = A_{\tilde{w}_{t+3}|w_{t+2},e^*}D_{e^*}A_{w_{t+2},w_{t+1}|e^*}
\] (B.12)

Consequently, \( Pr(\tilde{w}_{t+3}|w_{t+2}, e^*) \) is identified because \( \tilde{w}_{t+1} \) is a partition of \( w_{t+1} \). Moreover, the equilibrium selection mechanism can be identified using marginal distributions of \( Pr(\tilde{w}_{t+3}|w_{t+2}) \) through
\[
Pr(\tilde{w}_{t+3}|w_{t+2}) = \sum_{e^*} Pr(\tilde{w}_{t+3}|w_{t+2}, e^*) Pr(e^*)
\]

Note that \( \tilde{w}_{t+3} \) is a partition of \( w_{t+3} \), but our goal is to identify the Markov law of motion \( Pr(w_{t+3}|w_{t+2}, e^*) \), which satisfies the following equation.
\[
F_{w_{t+3},w_{t+2},\tilde{w}_{t+1}} = A_{w_{t+3}|w_{t+2},e^*}D_{e^*}A_{w_{t+2},\tilde{w}_{t+1}|e^*}
\] (B.13)

Since matrices \( D_{e^*} \) and \( A_{w_{t+2},\tilde{w}_{t+1}|e^*} \) are invertible, the Markov law of motion \( Pr(w_{t+3}|w_{t+2}, e^*) \) is identified. Since \( Pr(w_{t+3}|w_{t+2}, e^*) = Pr(a_{t+3}, x_{t+3}|a_{t+2}, x_{t+2}, e^*) \), the state transition matrix, which is allowed to be equilibrium specific, can be identified by summing over \( a_{t+3} \) in the Markov law of motion
\[
Pr(a_{t+3}, x_{t+3}|a_{t+2}, x_{t+2}, e^*) = \sum_{a_{t+3}} Pr(a_{t+3}, x_{t+3}|a_{t+2}, x_{t+2}, e^*)
\]

As a result, the policy function can be identified through the following equation:
\[
Pr(a_{t+3}, x_{t+3}|a_{t+2}, x_{t+2}, e^*) = Pr(a_{t+3}|x_{t+3}, a_{t+2}, e^*) Pr(x_{t+3}|a_{t+2}, x_{t+2}, e^*)
\]

\[ \blacksquare \]

C Asymptotic Properties of Estimators

This section gives the proofs of consistency and asymptotic normality of estimators. First I present the consistent estimation of the rank of a square matrix from a sequential testing. Thus the lower bound
of the number of equilibria or the number of equilibria itself can be consistently estimated. Given that
the number of equilibria is consistently estimated, the CCPs under each equilibrium are consistently
estimated, as are the structural parameters.

The structural parameters estimated through two steps are consistent and asymptotically normal
when the true number of equilibria is known. However, the number of equilibria is estimated through
a series of hypothesis tests, which serves as a model selection procedure. Consequently, the post-
model-selection sampling distributions are mixtures with properties that are very different from what
is conventionally assumed. To better understand the post-model-selection sampling distribution, one
must take the model selection step into account. Moreover, because there is only one correct model, the
sampling distribution of the estimated parameters can include estimates made from incorrect models
as well as the correct one. The sampling distribution of the structural parameters is a mixture of
two distributions, and such mixtures can depart dramatically from the distributions that conventional
statistical inference assumes. Post-model selection sampling distributions can be highly non-normal,
very complex, and with unknown finite sample properties even when the model responsible for the data
happens to be selected. There can be substantial bias in the regression estimates, and conventional
tests and confidence intervals are undertaken at some peril. The most effective solution is to have two
random samples from the population of interest: a training sample and a test sample. The training
sample is used to arrive at a preferred model. The test sample is used to estimate the parameters of the
chosen model and to apply statistical inference. For the test sample, the model is known in advance.
When there is one sample, an option is to randomly partition that sample into two subsets-split-sample
approach (see Berk, Brown, and Zhao (2010)).

The sequential testing provides us with a consistent estimator of the rank of a square matrix $F$ with
dimension of $J \times J$, if we carefully choose the significance level along with the sample size. Denote the
true matrix as $F$ with each element as a probability estimated by a simple frequency, and denote the
estimator of matrix $F$ as $\hat{F}$. To derive the limiting distribution for the determinant statistic, I denote
$f = \text{vec}(F)$ for the column vectorization of matrix $F$. Since each element of $F$ is estimated through a
simple frequency, then the limiting behavior of the estimator $\text{vec}(F)$ is characterized by:

$$\sqrt{N}(\text{vec}(\hat{F}) - \text{vec}(F)) \to_d N(0, \Sigma_F)$$

where $N$ is the sample size and $\Sigma_F$ is a $J^2 \times J^2$ covariance matrix. A matrix of non-full rank implies that
the determinant of the matrix equals to zero. In addition, the determinant of a matrix is a well-behaved
function of all the elements in the matrix. Let $\lambda$ denote the determinant of $F$ such that $\lambda = G(\text{vec}(F))$,
where $G$ denote the determinant function. Thus, the estimator of determinant $\hat{\lambda} = G(\text{vec}(\hat{F}))$. The delta method can be used to prove that the determinant statistic is asymptotically normal. Specifically, we have $\hat{\lambda} \simeq \lambda + \frac{dG}{df}(\text{vec}(\hat{F}) - \text{vec}(F))$. Under the null hypothesis that the determinant of matrix $F$ equals zero, i.e., $\lambda = 0$, the limiting distribution for the determinant statistics can be represented as:

$$\sqrt{N} \hat{\lambda} \rightarrow_d N(0, \frac{dG}{df} \Sigma_F \frac{dG}{df})$$

A sequence of tests is performed to estimate the rank of the matrix. This procedure is given as follows: the sequence of hypotheses $H_{l0} : \det(F^l) = 0$ is tested against the alternatives $H_{l1} : \det(F^l) \neq 0$ in decreasing order starting at $l = 1$. If no rejection occurs until $l = 1$, set $Q = 1$. Each hypothesis in this sequence is tested by a t-test where the error variance is always estimated from the overall model. More formally, we have,

$$\hat{\lambda} = \min \{ J : |\sqrt{N} \hat{\lambda}| \geq c_{JN}, 0 \leq J \leq (K + 1)^m \}$$

where $N$ represents the sample size. If the critical value associated with the significant level $\alpha$ is set to be a constant, then the rank estimated through this testing procedure is not consistent. The reason is that the testing procedure rejects the truth with probability $\alpha$ even as the sample size goes to infinity. To obtain consistent estimates for the rank, the critical value is chosen associated with a sample dependent significant level in a way that $\alpha_N$ goes to zero as the sample size $N$ goes to infinity but not faster than a given rate. Hosoya (1989) shows that if $\alpha_N$ goes to zero as the sample size $N$ goes to infinity and also $\lim_{N \rightarrow \infty} \frac{\ln \alpha_N}{N} = 0$, then the rank estimate provided by the sequential testing procedure will converge in probability to the true rank.

The estimation of CCPs are consistent and asymptotically normal whether the number of equilibria is already known or estimated through the sequence test described above. When the number of equilibria is known, firstly, the matrices used for eigenvalue-eigenvector decomposition are consistently estimated. Secondly, as in Andrew, Chu, and Lancaster (1993), eigenvalues and eigenvectors can be represented as a smooth function of the elements consisting of the matrix for the decomposition. Thus, the CCPs are consistently estimated and asymptotically normal. The same logic applies to the equilibrium selection mechanism. When the number of equilibria is estimated, CCPs are consistently estimated because when sample size goes to infinity, we always pick the right number of equilibria. Thus the asymptotic distribution is the same whether the number of equilibria is known or through estimation, see Hannan and Quinn (1979) and Pötscher (1991).

Consistency of the structural parameters follows naturally as the number of equilibria and CCPs
are consistently estimated. Note that estimation of the number of equilibria and CCPs is performed in each different market type $s_d$. After that, the estimation of the structural parameters needs to use variation from different markets via pooling all CCPs from different markets together. Consequently, the sample size for each market should increases with the same speed as the total sample size goes to infinity. Now I am going to derive the asymptotic distribution for the structural parameters.

With the number of equilibria known, assumptions needed for asymptotical properties of structural estimators are stated in the following:

A1: $\Theta$ is a compact set.

A2: the true value $\theta$ is in the interior of $\Theta$

A3: as $T \to \infty$, $W_M \to W_0$ a.s. where $W_0$ is a non-stochastic positive definite matrix.

A4: $\theta$ satisfies $[\sigma(\theta_0) - \Gamma(\sigma(\theta_0))]' W_0 [\sigma(\theta_0) - \Gamma(\sigma(\theta_0))] = 0$ implies that $\theta = \theta_0$

A5: the functions $\pi$ are twice continuously differentiable in $\theta$.

A6: the matrix $[\nabla \theta \Gamma(\sigma(\theta_0))]' W_0 [\nabla \theta \Gamma(\sigma(\theta_0))]$ is non-singular

Assumptions A1-A3, A5, and A6 are standard technical conditions to ensure the problem is well behaved. Assumption A4 ensures that the parameter vector is identified.

Note that the number of equilibria is determined through a sequence test, thus here I derive the asymptotical distribution of the post-selection estimators. Denote the true number of equilibria and $\hat{m}$ denote the first step estimator of $Q^0$, and the post-selection estimator of $\hat{\theta}_m$, then the asymptotical distribution of $\hat{\theta}_m$ can be represented as:

$$Pr\{m^\frac{1}{2}(\hat{\theta}_m - \theta_0)|Q^0\} = \sum_Q Pr\{m^\frac{1}{2}(\hat{\theta}_m - \theta_0), \hat{Q}_m = Q|Q^0\}$$

$$= Pr\{m^\frac{1}{2}(\hat{\theta}_m - \theta_0), \hat{Q}_m = Q^0|Q^0\} + Pr\{m^\frac{1}{2}(\hat{\theta}_m - \theta_0), \hat{Q}_m \neq Q^0|Q^0\}$$

$$= Pr\{m^\frac{1}{2}(\hat{\theta}_m - \theta_0)|\hat{Q}_m = Q^0, Q^0\} Pr\{\hat{Q}_m = Q^0|Q^0\} + Pr\{m^\frac{1}{2}(\hat{\theta}_m - \theta_0)|\hat{Q}_m \neq Q^0, Q^0\} Pr\{\hat{Q}_m \neq Q^0|Q^0\}$$

Since I use sequential test to obtain the estimator of $\hat{Q}_m$, the number of equilibria is consistently estimated, i.e. $Pr\{\hat{Q}_m = Q^0|Q^0\} = 1 - \alpha_m$ where $\alpha_m \to 0$ as $m \to \infty$. Equivalently, the probability of choosing the wrong number of equilibria goes to zero as the sample size goes to infinity, i.e., $\lim_{m \to \infty} Pr\{\hat{Q}_m \neq Q^0|Q^0\} = 0$. Consequently, the asymptotic distribution of the post-selection estimator $\hat{\theta}_m$ can be obtained by the asymptotic distribution of the conditional distribution of selecting the correct model. Note that if we use the same sample to do the test and the estimation, then the selection process and the post-selection estimator are not independent, i.e. $Pr\{m^\frac{1}{2}(\hat{\theta}_m - \theta_0)|\hat{Q}_m = Q^0, Q^0\} \neq Pr\{m^\frac{1}{2}(\hat{\theta}_m - \theta_0)|Q^0\}$. Asymptotically independent between selection and estimation re-
quires extra conditions, which is out of scope in this paper. In finite samples, selection and estimation are dependent because the estimation is based on the selection and sometimes the wrong number of equilibria might be chosen. As a result, the post-selection sample is different from the sample that we already know the true number of equilibria. However, if we have two independent samples where we can use one to do the test and the other one to do the estimation, then we have
\[ \Pr\{m^{\frac{1}{2}}(\hat{\theta}_m - \theta_0) | Q_m = Q^0, Q^0\} = \Pr\{m^{\frac{1}{2}}(\hat{\theta}_m - \theta_0) | Q^0\}. \]
Consequently, the asymptotic distribution of the post-selection estimator \( \hat{\theta}_m \) can be represented as:
\[
\lim_{m \to \infty} \Pr\{m^{\frac{1}{2}}(\hat{\theta}_m - \theta_0) | Q^0\} = \Pr\{m^{\frac{1}{2}}(\hat{\theta}_m - \theta_0) | Q^0\}
\]
where \( \Pr\{m^{\frac{1}{2}}(\hat{\theta}_m - \theta_0) | Q^0\} \) is asymptotically normal with regular conditions. Consequently, asymptotic normality is obtained using split-sample approach.

D Set Identification

There are some situations under which point identification is not feasible in the first step. For instance, the number of equilibria might not be point identified but instead identified up to a lower bound. This is because point identification requires that the possible actions or the number of players are relatively large compared to the number of equilibria. In a 2 × 2 entry game, using the methodology provided in this paper, one can only conclude that the number of equilibria is greater than 1 with the presence of multiplicity, i.e., the number of equilibria can only be identified to a lower bound. Consequently, all the other parameters of interest are not point identified. Moreover, using rank inequality to infer the number of equilibria needs a full rank condition on the unobserved matrix, which is empirically not testable. With this concern, partial identification is a possible direction to go. Note that no restriction is imposed to identify the lower bound, denoted as \( Q(s) \).

Suppose the payoff functions are known up to a vector parameter \( \theta \). The mixture feature can be represented as:
\[ Pr(a|s, \theta) = \sum_{e^*} \prod_i Pr(a_i|s, \theta, e^*) Pr(e^*), \text{ with } \sum_{e^*} Pr(e^*) = 1. \]
Since the economic theory does not predict which equilibrium players tend to choose, no restriction is imposed in the equilibrium selection mechanism. Specifically, the equilibrium selection mechanism is allowed to be any valid mixture across equilibria, i.e., \( \sum_{e^*} Pr(e^*) = 1 \) and \( Pr(e^*) \geq 0 \). The identified set can be represented
as:
\[
\Theta_I = \left\{ \theta \in \Theta : \forall a, s, Pr(e^*) \in [0, 1], s.t. \ Pr(a|s, \theta) = \sum_{e^*} \prod_i Pr(a_i|s, \theta, e^*) Pr(e^*) \right\}
\]

where \(Pr(a_i|s, \theta, e^*)\) is CCPs under each equilibrium, and \(\varepsilon(s, \theta)\) is the collection of all the equilibria. The intuition of deriving the identified set is that for a given value of \(\theta\), one can solve for all the equilibria \(Pr(a_i|s, \theta, e^*)\) via Homotopy methods conditional on \(s\). For this \(\theta\) that rationalizes the observed outcome, there must exist a valid equilibrium selection mechanism that satisfied the above equation. This has to be satisfied for all possible values of market characteristics.

Note that the identified set is sharp because I employ the lower bound of the number of equilibria as a screen of the structural parameter \(\theta\). Those \(\theta\) that yields a number of equilibria smaller than the lower bound are excluded in the identified set. In the above 2 \(\times\) 2 entry game example, if the number of equilibria is 1, then point identification is obtained; if the number of equilibria is greater than 1, then those \(\theta\) which provides a singleton equilibrium are deleted from the identified set.

E Graphs and Tables

Figure 1: The Best Response Function for Different Values of \(s\)
Table 1: Monte Carlo Evidence: Cross-sectional Data

<table>
<thead>
<tr>
<th>The Number of Equilibria</th>
<th>DGP Estimates</th>
<th>DGP Estimates</th>
<th>DGP Estimates</th>
<th>DGP Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma(a = 0</td>
<td>s, e^* = 1)$</td>
<td>0.1593 (0.012)</td>
<td>0.1780 (0.014)</td>
<td>0.2053 (0.012)</td>
</tr>
<tr>
<td>$\sigma(a = 0</td>
<td>s, e^* = 2)$</td>
<td>0.8671 (0.013)</td>
<td>0.8772 (0.012)</td>
<td>0.8859 (0.013)</td>
</tr>
<tr>
<td>The Equilibrium Selection</td>
<td>0.5 (0.017)</td>
<td>0.5 (0.018)</td>
<td>0.5 (0.018)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

1 The number in brackets is the standard deviation computed through bootstrap with 500 repetition
2 Sample size of each market type is 1200

Table 2: Monte Carlo Evidence: Model Primitives

<table>
<thead>
<tr>
<th>DGP</th>
<th>Cross-section</th>
<th>Panel data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unique Eq</td>
<td>Multiple Eq</td>
</tr>
<tr>
<td>Strategic Interaction $\beta$</td>
<td>2.5 (0.0405)</td>
<td>2.5054 (0.0190)</td>
</tr>
<tr>
<td>Market Effect $\alpha$</td>
<td>-0.0236 (0.0056)</td>
<td>0.0398 (0.0048)</td>
</tr>
</tbody>
</table>

1 The number in brackets is the standard deviation computed through bootstrap with 500 repetition
2 Sample size of each market type is 1200
Figure 3: Timing Patterns for Commercials across Markets

Table 3: Monte Carlo Evidence: Panel Data

<table>
<thead>
<tr>
<th>s=1</th>
<th>s=2</th>
<th>s=3</th>
<th>s=4</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGP</td>
<td>Estimates</td>
<td>DGP</td>
<td>Estimates</td>
</tr>
<tr>
<td>σ(α = 0</td>
<td>s, e∗ = 1)</td>
<td>0.1593</td>
<td>0.1605</td>
</tr>
<tr>
<td>(0.017)</td>
<td>(0.018)</td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>σ(α = 0</td>
<td>s, e∗ = 2)</td>
<td>0.8671</td>
<td>0.8581</td>
</tr>
<tr>
<td>(0.0322)</td>
<td>(0.017)</td>
<td>(0.028)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>prob(e∗ = 2</td>
<td>e∗ = 1)</td>
<td>0.1</td>
<td>0.094</td>
</tr>
<tr>
<td>(0.027)</td>
<td>(0.045)</td>
<td></td>
<td>(0.042)</td>
</tr>
<tr>
<td>prob(e∗ = 2</td>
<td>e∗ = 2)</td>
<td>0.8</td>
<td>0.826</td>
</tr>
<tr>
<td>(0.050)</td>
<td>(0.031)</td>
<td></td>
<td>(0.046)</td>
</tr>
</tbody>
</table>

1 The number in brackets is the standard deviation computed through bootstrap with 500 repetition
2 Sample size of each market type is 1200

Table 4: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>std. Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. Players</td>
<td>108554</td>
<td>12.93453</td>
<td>3.174468</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Timing</td>
<td>108554</td>
<td>.4985537</td>
<td>.5000002</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Day</td>
<td>108554</td>
<td>31.42016</td>
<td>17.56653</td>
<td>1</td>
<td>59</td>
</tr>
<tr>
<td>Hour</td>
<td>108554</td>
<td>16.46269</td>
<td>3.153961</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>Market(big=1)</td>
<td>108554</td>
<td>.5168672</td>
<td>.4997177</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 4: Timing Patterns for Commercials in Different Markets

TIMING OF COMMERCIALS IN ORLANDO, FL., AND ROCHESTER, N.Y., ON OCTOBER 30, 2001 5-6 P.M.

(A) Orlando, FL.

(B) Rochester, N.Y.
### Table 5: Estimation of Commercial Airing Strategies with Cross-sectional Data

<table>
<thead>
<tr>
<th></th>
<th>All market</th>
<th>Large</th>
<th>Small</th>
<th>Drive-time</th>
<th>Non drive-time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Market Size</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4-5 PM</td>
<td>5-6 PM</td>
<td>12-1 PM</td>
<td>9-10PM</td>
<td></td>
</tr>
<tr>
<td><strong>The Number of Equilibria</strong></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>The Equilibrium Selection:</strong> $\text{prob}(e^* = 1)$</td>
<td>0.2846</td>
<td>1.0000</td>
<td>0.2732</td>
<td>0.3307</td>
<td>0.4192</td>
</tr>
<tr>
<td></td>
<td>(0.1293)</td>
<td>(0.0192)</td>
<td>(0.1191)</td>
<td>(0.1275)</td>
<td>-</td>
</tr>
<tr>
<td>**CCPs: $\text{prob}(a = 0</td>
<td>e^* = 1)$**</td>
<td>0.6565</td>
<td>0.6582</td>
<td>0.6841</td>
<td>0.6687</td>
</tr>
<tr>
<td></td>
<td>(0.1875)</td>
<td>(0.0999)</td>
<td>(0.1171)</td>
<td>(0.2206)</td>
<td>(0.0043)</td>
</tr>
<tr>
<td>**$\text{prob}(a = 0</td>
<td>e^* = 2)$**</td>
<td>0.4288</td>
<td>0.4287</td>
<td>0.3974</td>
<td>0.3700</td>
</tr>
<tr>
<td></td>
<td>(0.0630)</td>
<td>(0.0498)</td>
<td>(0.0884)</td>
<td>(0.1107)</td>
<td>-</td>
</tr>
<tr>
<td><strong>$\alpha$</strong></td>
<td>-0.0055</td>
<td>-0.0242</td>
<td>-0.0056</td>
<td>-0.0092</td>
<td>-0.0051</td>
</tr>
<tr>
<td></td>
<td>(0.3258)</td>
<td>(0.0389)</td>
<td>(0.1007)</td>
<td>(0.3019)</td>
<td>(0.0043)</td>
</tr>
<tr>
<td><strong>$\beta$</strong></td>
<td>2.0520</td>
<td>0.20532</td>
<td>2.0736</td>
<td>2.0665</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0.3147)</td>
<td>(0.1089)</td>
<td>(0.2844)</td>
<td>(0.1999)</td>
<td>-</td>
</tr>
</tbody>
</table>

1 The number in brackets is the standard deviation computed through bootstrap with 500 repetition.
2 For markets with unique equilibrium, $\delta$ is assumed to be zero.

### Table 6: Estimation of Commercial Airing Strategies with Panel Data

<table>
<thead>
<tr>
<th></th>
<th>All Market</th>
<th>Large</th>
<th>Small</th>
<th>Drive-time</th>
<th>Non drive-time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Market Size</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4-5 PM</td>
<td>5-6 PM</td>
<td>12-1 PM</td>
<td>9-10PM</td>
<td></td>
</tr>
<tr>
<td><strong>The Number of Equilibria</strong></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Eq Evolution:</strong> $\text{pr}(e_{t+1}^* = 1</td>
<td>e_t^* = 1)$</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>(0.0192)</td>
<td>(0.4311)</td>
<td>(0.0814)</td>
<td>(0.0485)</td>
<td>-</td>
</tr>
<tr>
<td>*<em>$\text{pr}(e_{t+1}^</em> = 2</td>
<td>e_t^* = 2)$**</td>
<td>0.8714</td>
<td>-</td>
<td>0.9777</td>
<td>0.8749</td>
</tr>
<tr>
<td></td>
<td>(0.0192)</td>
<td>(0.4076)</td>
<td>(0.0959)</td>
<td>(0.0908)</td>
<td>-</td>
</tr>
<tr>
<td>**CCPs: $\text{prob}(a = 0</td>
<td>e^* = 1)$**</td>
<td>0.5612</td>
<td>0.5116</td>
<td>0.6576</td>
<td>0.5769</td>
</tr>
<tr>
<td></td>
<td>(0.0234)</td>
<td>(0.0085)</td>
<td>(0.1523)</td>
<td>(0.0384)</td>
<td>(0.0414)</td>
</tr>
<tr>
<td>**$\text{prob}(a = 0</td>
<td>e^* = 2)$**</td>
<td>0.3617</td>
<td>-</td>
<td>0.4382</td>
<td>0.3037</td>
</tr>
<tr>
<td></td>
<td>(0.0371)</td>
<td>(0.0608)</td>
<td>(0.0719)</td>
<td>(0.0304)</td>
<td>-</td>
</tr>
<tr>
<td><strong>$\alpha$</strong></td>
<td>0.0037</td>
<td>-0.0465</td>
<td>-0.0053</td>
<td>0.0108</td>
<td>-0.0022</td>
</tr>
<tr>
<td></td>
<td>(0.0152)</td>
<td>(0.0340)</td>
<td>(0.0622)</td>
<td>(0.0195)</td>
<td>(0.0301)</td>
</tr>
<tr>
<td><strong>$\beta$</strong></td>
<td>2.0402</td>
<td>0.20055</td>
<td>2.0860</td>
<td>2.0500</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0.0336)</td>
<td>(0.0719)</td>
<td>(0.2401)</td>
<td>(0.0335)</td>
<td>-</td>
</tr>
</tbody>
</table>

1 The number in brackets is the standard deviation computed through bootstrap with replacement, with 500 repetition.
2 For markets with unique equilibrium, $\delta$ is assumed to be zero.