IDENTIFICATION OF INCOMPLETE INFORMATION GAMES WITH MULTIPLE EQUILIBRIA AND UNOBSERVED HETEROGENEITY

by

Ruli Xiao

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

June, 2014

Abstract

This dissertation mainly studies identification of finite action games with incomplete information. The essential contribution of this dissertation is to allow for the presence of multiple equilibria and/or unobserved market-level heterogeneity. Chapter 2 provides a novel methodology to nonparametrically identify static games with multiple equilibria. Exploiting the results in mis-classification error models, I show that the number of equilibria, the equilibrium selection mechanism and individual equilibrium strategies associated with all positively employed equilibria can be nonparametrically identified from the distributions of the game outcomes. Provide the equilibrium conditional choice probabilities, payoffs then can be identified nonparametrically with exclusion restrictions. A natural estimator is also proposed following the constructive identification procedure. The empirical application investigates the strategic interaction among radio stations when they choose commercial timings, which provides evidence that two equilibria exist.

Chapter 3 extends chapter 2 to incorporate unobserved market-level heterogeneity. This chapter assumes that the market-level latent type is discrete and has a finite support. With the discrete feature, the presence of unobserved heterogeneity generates similar finite mixture feature as the presence of multiple equilibria. The combination of both payoff-

relevant and payoff un-relevant latent factor complicates the identification because of lacking information to disentangle the two. Consequently, instead of providing point identification, I provide set identification for the payoff parameters in chapter 3. To understand the trade-off between point identification and extra assumptions, I also provide conditions under which the identified set shrinks to a point.

Chapter 4 considers identification in dynamic settings. If only Markov Perfect Equilibria being considered, observables including actions and payoff relevant covariates in period t follow a first-order Markov process in time series by a market. This Markov property is a key condition under which dynamic games can be nonparametric identified with four periods of data. In particular, the law of motion associated with every possible combination of equilibria and the unobserved market-types can be nonparametrically identified. Additionally, payoffs can be identified nonparametrically with exclusion restrictions. More importantly, multiple equilibria and unobserved heterogeneity can be distinguished from the test with the null that payoffs associated with two levels of latent factor are the same. Specifically, if two payoffs are the same, then they should belong to the same latent market type but different equilibria. On the other hand, if two payoffs are different, they should be driven by the heterogeneity.

Chapter 5 concludes and proposes possible avenues for future research based on this dissertation.

Keywords:	Incomplete Information, Multiple Equilibria, Unobserved heterogeneity
	Nonparametric Identification, Partial Identification
	Static Game, Dynamic Game, Measurement Error Models
JEL Classification:	C14, D03, D44
Advisors:	Professor Yingyao Hu
	Professor Yuya Sasaki

Acknowledgements

First, this thesis would not exist without the patient ear and expert guidance of my advisor, Professor Yingyao Hu. I am deeply grateful for his guidance, understanding and support during my graduate studies at Johns Hopkins University.

I am also deeply indebted to Professor Yuya Sasaki for his guidance and support, to Professor Hulya Eraslan for her valuable comments, to Professor Andrew Sweeting for the data and suggestions, to Professor Yonghong An for his comments and moral support. I thank all graduates students at Hopkins, especially Natsuki Arai, Ryan Bush, Mikhail Smirnov, Sangah Li, Lucia Tiererova, Gwyn Pauley, Michele Mazzoleni, Xi Yang and Hou Wang, who come to my presentation, help revising my drafts and be supportive.

I could not thank more to my parents and my parents-in-law, who come here from China just to help. I would not manage without their support and tolerance. Thank you also to my sister's family and my brother's family.

Finally, I dedicate this dissertation to my husband, Dr. Zhen Li and my adorable little boy Adam. Life wouldn't be complete without them.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

During the past decade, estimating empirical models of games has become an important and active research area in industrial organization, applied econometrics and marketing. See Seim (2006), Sweeting (2011), Bajari, Benkard, and Levin (2007) as well as Aguirregabiria and Mira (2007). See also Nevo and Aguirregabiria (2010) for survey. Multiple equilibria and unobserved heterogeneity are two outstanding and important features in game analysis. In viewing of these, this dissertation contributes to the literature mainly by providing a novel identification methodology for finite action games while allowing for both multiplicity of equilibria and/or market-level discrete unobserved heterogeneity. The game studied here is assumed to have incomplete information that the incompleteness due to the fact that the payoff shocks are only observable to the individual player herself but not to her rivals.

Chapter 2 of this dissertation tackles the problem of multiple equilibria in static game settings. It provides a nonparametric analysis of finite action games with incomplete information while does not impose any ad hoc assumptions on the equilibrium selection or the uniqueness of the equilibrium. The identification uses results from measurement error literature by treating the underlying equilibrium as a latent variable. Thus, I prove that all aspects of the game such as the number of equilibria, the equilibrium selection mechanism, the equilibrium strategies of individual players associated with each equilibrium are nonparametrically identified. Moreover, the payoff functions are nonparametrically identified with exclusion restrictions.

To apply the identification technique developed by Hu (2008), one essential condition is to find measures or proxies for the latent equilibria, and those measures are independent conditional on the equilibrium employment. I consider identification in both cross-sectional and panel data scenarios. If cross-section data is available, the traditional assumption that private payoff shocks are independent across actions and players plays an important role in identification. With this assumption holds, individual players' actions can serve as measures for the underlying equilibria. As a result, the observed joint distributions of players can be expressed as a mixture over equilibrium choice probabilities of individual players with the equilibrium selection probability as the mixture weight. Therefore, the number of equilibria can be identified as the rank of the matrix constructed by joint distributions of actions of two players, provided with enough number of alternatives and full rank condition satisfied. Moreover, the equilibrium choice probabilities can be identified as eigenvalues of the matrix estimated from the data directly. Finally, the payoff can be identified with exclusion restrictions following the existing literature.

When panel data is available, I provide identification allowing for the equilibrium employment evolves according to a first-order Markov process instead of assuming that the same equilibrium is employed over time, as the literature tends to assume. With the static framework, the only factor that generates correlation between actions over time is the underlying equilibrium evolution. Otherwise actions over time are independent. Thus, actions in previous period and next period can serve as the different two measures for the underlying equilibrium employed in this period. If we can get access to three periods of data, the equilibrium relevant aspects of the game therefore are nonparametrically identified. Identification of payoff is exactly the same as the case with cross-sectional data.

Given that the identification is constructive, estimation follows naturally. As an application, I take the estimation procedure to the field data where radio stations strategically determine when to air their commercials. I find that there are two equilibria exist in smaller markets while unique equilibrium exist in large market. Among smaller markets, the two equilibria are employed with a probability of around 0.7 versus 0.3 respectively. This unequal selection probability is worth noting, and providing rational might trace back to different culture or preference. Moreover, I find that markets employ the same equilibrium over time exploring the panel data feature. This provides empirical evidence, and we might be more comfortable to assume that the same equilibrium is employed over time in different settings.

Chapter 2 obtains identification for static games with incomplete information allowing for multiple equilibria while assumes away market level unobserved heterogeneity. However, existence of discrete unobserved market-level heterogeneity generates the same finite mixture feature as with multiple equilibria. Furthermore, incorporate unobserved heterogeneity is important in empirical studies not only in games but also in discrete choice models. It is highly possible that econometrician cannot gather all relevant information. Thus, controlling for unobserved factors is important for recovering the strategic interaction between players. Consequently, Chapter 3 extends chapter 2 to incorporate unobserved market-level types into the identification.

Existence of multiple equilibria and unobserved market heterogeneity result in the same mixture structure of the observed joint distributions. In alignment with this spirit, I create a new latent variable that combines information of equilibria and heterogeneity. Similar as Chapter 2, the cardinality of this new variable is identified as the rank of the matrix constructed by the joint distribution of individual players. Also choice probabilities associated with this latent variable are identified up to ordering. The difficulty of the identification lies in the requirement to order the latent variable. From the matrix eigenvalue-eigenvector decomposition, we show that the choice probabilities can be identified as eigenvalues without ordering. The key different between chapter 2 and chapter 3 is that ordering of equilibria does not have economic meanings. Thus, it is trivial, and any ordering is fine. In contrast, ordering of market heterogeneity is important because it reflects how payoff differs in different market types. Disentangling the two unobserved factors is possible only when for some observables and all unobserved market types, a unique equilibrium is guaranteed. Instead of point identification, this paper focuses on set identification which considers all possible ordering.

Existing literature have focused on static environments or on single-agent dynamic decision problems. Many economic policy debates, however, turn on quantities that are inherently linked to dynamic competition, such as entry and exit costs, the returns to advertising or research and development, the adjustment costs of investment, or the speed of firm and consumer learning. Estimating these dynamic parameters has been seen as a major challenge, both conceptually and computationally. For estimation purpose, two-step estimation approaches pioneered by Hotz and Miller (1993) are prevalent in applications. To make estimation tractable, existing literature usually assume that the data is generated by the same equilibrium. Moreover, even though unobserved heterogeneity sometimes are controlled for in estimation, for instance Aguirregabiria and Mira (2007), there is no identification results on dynamic games with unobserved heterogeneity. To address both issues, Chapter 4 provides identification and estimation for dynamic games with incomplete information while allowing for both multiplicity of equilibria and finite unobserved market heterogeneity.

Chapter 4 imposes no restrictions on the cardinality of the equilibrium set or the equilibrium selection rules in identification. Identification proceeds in the following steps. First I identify the cardinality of a new latent variable which combines information of both the unobserved market-type and the multiple equilibria. Second I identify the law of transition for the Markov process, thus the equilibrium specific conditional choice probability and the transition function for both observed and unobserved state variables are identified. Third, the payoff function are nonparametrically identified with exclusion restrictions as in Pesendorfer and Schmidt-Dengler (2008) for each value of the new latent factor. Consequently, one can distinguish between multiple equilibria and unobserved-market types from comparing the payoff functions. Specifically, multiple equilibria map with the same payoff functions while unobserved-market types are associated with different level of payoffs. As a byproduct, the equilibrium selection and the marginal distribution of the market-type can be identified.

In Chapter 5, I conclude and summarize the dissertation. Furthermore, I also present several avenues for future research based on this dissertation.

Chapter 2

Identification and Estimation of Static Games with Multiple Equilibria

2.1 Introduction

Unlike single-agent discrete choice models, games generally admit multiple equilibria. Although multiplicity in games does not necessarily preclude estimation, ignoring it might result in mis-specification. Moreover, even if the game primitives can be consistently estimated under some assumptions (e.g., a unique equilibrium assumption in the data as in Bajari, Hong, Krainer, and Nekipelov (2010b) and Aradillas-Lopez (2010)), it is 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[1]. 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, which allows indexing the equilibria. With 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

---

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

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 implemented and computationally convenient method to estimate game primitives. As is widely know, 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 static game with both cross-sectional and panel data structures. When we only get access to 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 at least three players in the game. 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. Correlation between actions of all players across time is the identification power.

Applying the proposed methodology to the data on radio stations studied in Sweeting (2009), 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 Sweeting (2009). 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. I begin with a literature review in section 2. Then section 3 outlines the static game framework, and provides the nonparametric identification of the game. Next section 4 describes the estimation procedure for static games. After that section 5 provides a Monte Carlo illustration to provide evidence for the methodology. Then section 6 provides an empirical application to radio station commercial airing. Lastly, section 8 concludes. The Appendix contains the proofs, the figures and the tables.

2.2 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 the game primitives, it is nearly impossible to simulate the model and provide counterfactual inference. Moreover, the unique equilibrium assumption is lack of 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 computed, one approach is to focus on certain quantities are invariant across 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 different equilibria. Tamer (2003) nevertheless identifies the game primitives using exclusion restrictions and large support conditions on the observable covariates. The identification results rely on the extreme values of covariates by reducing the problem to a single-agent decision problem. However, this identification-at-infinity strategy leads to a slower asymptotic convergence rate (see 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[2]. Bajari, Hong, and

[2]Bounds estimation has also been used by Ciliberto and Tamer (2009), Pakes, Porter, Ho, and Ishii

Ryan (2010) incorporate a parameterized equilibrium selection function into the problem[3], 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. The paper is also related to Aguirregabiria and Mira (2013), which focuses on distinguish of multiple equilibria and payoff relevant heterogeneity.

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.

2.3 Static Game Setup

Consider a static simultaneous move game that involves N players. Players obtain action specific 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, $i \in \{1, ..., N\}$, chooses an action a_i out of a finite set $\mathscr{A} = \{0, 1, ..., K\}$. Let a_{-i} denote player i's rivals'

(2006), and Andrews, Berry, and Jia (2004). Berry and Tamer (2006) and Berry and Reiss (2007) survey the econometric analysis of discrete games.

[3] See also Ackerberg and Gowrisankaran (2006) and Bjorn and Vuong (1984).

actions and $x \in \mathscr{X}$ denote public observable state variable. The $K+1$ action specific profit shocks are denoted as $\epsilon_i(a_i)$, and their density distributions are denoted as $f(\epsilon_i)$[4]. The payoff for player i from choosing action a_i is assumed to be additive separable as below:

$$U_i(a_i, a_{-i}, x, \epsilon_i) = \pi_i(a_i, a_{-i}, x) + \epsilon_i(a_i)$$

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

Instead of defining the equilibrium using players' decision rules, I defined the equilibrium using the probability that each player choosing each possible action, i.e. $\sigma_i(a_i|x)$ denotes the probability that player i chooses action a_i conditional on observing x. Since player i does not observe her rivals' payoff shocks, she has to form belief over the distribution of actions that her rivals are going to choose. Meaning, player i needs to best response to her rivals' action distribution instead of a specific action. With the following independence assumption, player i's belief does not depend on her own private information.

Assumption 2.1. *(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.*

The assumption of conditional independence among 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),

[4]Similar setups are studied in Seim (2006) and 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. Consequently, player i expects her rivals' action distribution to be $\sigma_{-i}(a_{-i}|x)$. Then I can represent player i's expected utility from choosing action a_i as the following:

$$u_i(a_i, x, \epsilon_i) = \sum_{a_{-i}} \pi_i(a_i, a_{-i}, x)\sigma_{-i}(a_{-i}|x) + \epsilon_i(a_i) \equiv \Pi_i(a_i, x) + \epsilon_i(a_i)$$

The Bayesian Nash Equilibrium is stated in the following:

Definition 2.2. ***(BNE)*** *For a fixed state x, the Bayesian Nash Equilibrium (BNE) is a collection of probabilities $\sigma_i^*(a_i = k|x)$ for $i = 1, ..., N$ and $k = 0, ..., K$ such that for all i and k, the following equation satisfied:*

$$\sigma_i(a_i = k|x) = \Pr\left(\Pi_i(a_i = k, x) + \epsilon_i(a_i = k) > \Pi_i(a_i = j, x) + \epsilon_i(a_i = j), \forall j\right)$$

Following Hotz and Miller (1993), the equilibrium condition implies a one-to-one mapping between the CCPs ($\{\sigma_1^*(a_1|x), ..., \sigma_n^*(a_n|x)\}$)and the difference of expected choice utilities ($\{\Pi_1(a_1, x) - \Pi_1(a_1 = 0, x), ..., \Pi_n(a_n, x) - \Pi_n(a_n = 0, x)\}$)

Assumption 1 guarantees that the mapping has at least one fixed point by Brouwer's fixed point theorem. As a result, if CCPs are obtained, the differences of the expected choice values are as well. If there is a unique equilibrium, CCPs computed from the collected data can be used to approximate the CCPs predicted by theory. Then 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 players favor one equilibrium over the other. 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.

2.4 Nonparametric Identification Results

Before I go into detail about the identification, let me first describe the data structure. Suppose the econometrician observes actions of all players in cross-sectional markets m where $m = 1, ..., M$ with characteristics $s_m \in S$. Assume that S is discrete and has a finite support. Assume that the number of equilibria is finite, I can index the equilibrium by $e^* \in \Omega_{\pi,x} \equiv \{1, ..., Q_x\}$, where Q_x is the number of equilibria[5]. Note that the cardinality of the equilibrium set varies with game characteristics. Identification of equilibrium-specific components is established conditional on the market characteristics x_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.

$$\Pr(a_1, ..., a_n) = \sum_{e^*} \Pr(a_1, ..., a_n|e^*) \Pr(e^*) = \sum_{e^*} \prod_i \Pr(a_i|e^*) \Pr(e^*) \tag{2.1}$$

where $\Pr(e^*)$ is the probability that equilibrium e^* is employed, i.e., the equilibrium selection mechanism, and $\Pr(a_i|e^*)$ is the CCPs of player i associated with equilibrium e^*. The first equality is due to the law of total probability while the second equality hold because the independent assumption of private information.

If the data is generated by the same equilibrium, the joint distributions of players' actions from the data forms an equilibrium joint distribution. Thus, one can infer whether the data is generated by the same equilibrium by testing whether players' actions are independent. Failing to reject the null hypothesis that this condition holds, one can argue that there is unique equilibrium in the data 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 the identification. Before I move to the identification detail, let me first define what identification means.

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

[5]If the number of equilibria is infinite, it is impossible to do the identification using limited information. Also by (Harsanyi (1973)), we know that the game that has infinite number of equilibria has zero measure

if there does not exist another set of $\{Q, p, P\}$ that is consistent with the observed data $\{a_1^m, ..., a_n^m, x_m\}_m$.

Identification of the number of equilibria can be obtained using information of correlation between two representative players. With two players, the joint distribution becomes:

$$\Pr(a_1, a_2) = \sum_{e^*} \Pr(a_1|e^*) \Pr(a_2|e^*) \Pr(e^*) \tag{2.2}$$

Matrixes are used extensively in this paper to make use of all possible information together at the same time. I rewrite this equation into matrix form:

$$F_{a_1,a_2} = A_{a_1|e^*} D_{e^*} A^T_{a_2|e^*} \tag{2.3}$$

where

$$\begin{aligned} F_{a_1,a_2} &\equiv [\Pr(a_1 = k, a_2 = j)]_{k,j}, \\ A_{a_i|e^*} &\equiv [\Pr(a_i = k|e^* = q)]_{k,q} \\ D_{e^*} &\equiv \operatorname{diag}\Big[\Pr(e^* = 1) \ \ ... \ \ \Pr(e^* = Q)\Big]. \end{aligned}$$

Those matrices stack the distributions with all possible values that a_1, a_2 and e^* can take. In particular, matrix F_{a_1,a_2} consists of the whole joint distributions of a_1 and a_2, which can be estimated from data. D_{e^*} is a diagonal matrix with the probability of each equilibrium being selected as the diagonal elements, while matrix $A_{a_i|e^*}$ collects all the CCPs associated with equilibrium e^*.

The dimensions of the three matrices defined above F_{a_1,a_2}, $A_{a_i|e^*}$ and D_{e^*} 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 2.3. *The rank of the observed matrix F_{a_1,a_2} serves as the lower bound for the number of equilibria, i.e., $Q \geq Rank(F_{a_1,a_2})$. Furthermore, the number of equilibria is identified, particularly, $Q = Rank(F_{a_1,a_2})$ if the following conditions are satisfied:*
(1) $K+1 > Q$; (2) both matrices $A_{a_1|e^*}$ and $A_{a_2|e^*}$ are full rank; (3) all $\Pr(e^*)$ are positive.

Proof See Appendix. □

The first condition requires that the number of possible actions 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. However, here I only use data from two players. If there are more players, grouping all players into two groups expands the action space that a representative player can choose from. Consequently, it increases identification power.

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. 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 in the data 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 the number of equilibria active in the data. Note that here I only use information from two players to recover the number of equilibria and the first condition seems is restrictive because it does not hold in the 2×2 game framework. However, if there are more than two players, we can use information from more players by grouping them into two groups. This treatment will expand the dimension of the corresponding matrix and increase identification power.

Provided that the number of equilibria is identified, I show below how to identify CCPs associated with different equilibria for each individual player. In measurement error literature, normally the latent variable share the same support with its measurements. Thus, the full rank condition indicates that the relevant matrices are invertible. In order to follow the technique used in measurement error literature which uses matrix eigenvalue-eigenvector decomposition, first I need to tailor matrices to be square. To do that, I partition the action space so that it has the same dimension as the number of equilibria. With a little bit abuse of notation, I still use $\tilde{a}_1$ and $\tilde{a}_2$ to represent the action variable after the domain partition. One key criteria for the partition is that the corresponding trailed matrix $A_{\tilde{a}_i|e^*}$ is full rank for both $i = 1, 2$. The following lemma states that there exists at least one way to achieve the partition.

Lemma 2.4. *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 $\square$

Identification requires extra information, for example, another player which I call as player 3. Then the following equation links observed joint distribution with those unknowns still holds:

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

To use the matrix algebra for identification, I fix $a_3 = k$ and use all possible variation from player 1 and 2's actions. Matrices definition is the same as I defined above with the only difference is that a_3 is fixed. Using the above matrix representations, we have following two equations:

$$F_{\tilde{a}_1,\tilde{a}_2} = A_{\tilde{a}_1|e^*} D_{e^*} A^T_{\tilde{a}_2|e^*} \tag{2.4}$$

$$F_{\tilde{a}_1,\tilde{a}_2,a_3=k} = A_{\tilde{a}_1|e^*} D_{a_3=k|e^*} D_{e^*} A^T_{\tilde{a}_2|e^*} \tag{2.5}$$

Given that $A_{\tilde{a}_1|e^*}$ and $A_{\tilde{a}_2|e^*}$ have full rank, post-multiplying $F^{-1}_{\tilde{a}_1,\tilde{a}_2}$ on both sides of equation

2.5 leads to the following main equation, which is essential for the identification.

$$F_{\tilde{a}_1,\tilde{a}_2,a_3=k}F^{-1}_{\tilde{a}_1,\tilde{a}_2} = A_{\tilde{a}_1|e^*}D_{a_3=k|e^*}A^{-1}_{\tilde{a}_1|e^*} \tag{2.6}$$

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 2.5. *(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 2.5, the eigenvalue-eigenvector decomposition in equation 2.6 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 2.6. *With assumptions 1 and 2, and conditions in lemma 1 satisfied, CCPs of players in each equilibrium and the equilibrium selection are nonparametrically identified.*

Proof See Appendix B $\square$

Based on lemmas 1, 2, and 3, conditional on market characteristic x, I have already identified all the CCPs under every employed equilibrium $\Pr(a_i|x, e^* = 1)\ldots\ldots \Pr(a_i|x, e^* = Q_x), \forall i, a_i$. Each equilibrium $\Pr(a_i|x, e^* = q)$ 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.

As with analysis in discrete choice models, it is impossible to separately identify all the payoff functions, only their differences can be recovered. Normalization is necessary and stated in the following assumption:

Assumption 2.7. ***(Normalization)*** *For all i and all a_{-i} and x, $\pi_i(a_i = 0, a_{-i}, x) = 0$.*

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 and the extreme value distribution of shocks, the equilibrium condition becomes

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

Identification of payoff functions $\pi_i(a_i = k, a_{-i}, x)$ 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 2.8. ***(Exclusion Restriction)*** *For each player i, the state variable can be partitioned into two parts denoted as x_i, x_{-i}, so that only x_i enters player i's payoff, i.e. $\pi_i(a_i = k, a_{-i}, x) \equiv \pi_i(a_i = k, a_{-i}, x_i)$.*

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

$$log\sigma_i(a_i = k|x_i, x_{-i}, e^* = q) - log\sigma_i(a_i = 0|x_i, x_{-i}, e^* = q) = \sum_{a_{-i}} \pi_i(a_i = k, a_{-i}, x_i)\sigma_{-i}(a_{-i}|x_i, x_{-i}, e^* = q$$

Variation of x_{-i} expands the total number of equations without adding more unknowns, which helps for identifying payoff functions $\pi_i(a_i = k, a_{-i}, x_i)$ 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 x, there are $K \times Q_x$ equations, which is magnified by the number of equilibria Q_x, 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 ($Q_x \geq (K+1)^{n-1}$). Unfortunately, first step identification requires the number of multiple equilibria (Q_x) 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 2.9. ***(Identification of static game)*** *With assumptions 2.1, 2.5, 2.7, and 4.11 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.*

2.4.1 Nonparametric Identification with Panel Data

This subsection presents identification of static games with panel data structure. When one can get access to panel data, estimation can be done along time series on individual markets, which implicitly 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. Moreover, even if it is true that players employ the same equilibrium over time, another consideration is that estimation requires the same market is observed over a long period of time. Last, without knowing the equilibrium selection mechanism, it is impossible to conduct counterfactual analysis. In a word, it is important to still identify the underlying game structure without imposing the assumption that the same equilibrium is employed over time.

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., $a_t = \left(a_{1t}, ..., a_{nt} \right) \in \mathscr{A}^n$ where $\mathscr{A}^n$ 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_t^* denote the index of equilibria employed in time t, I assume that e_t^* follows a first-order

Markov process. This paper is not going to model how players manage to coordinate with choosing different equilibria over time.

Assumption 2.10. ***(First-order Markov Equilibrium Evolution)*** *The equilibrium that a typical market employs follows a first-order Markov process, i.e.,* $\Pr(e^*_{t+1}|x, e^*_t, ..., e^*_0) = \Pr(e^*_{t+1}|x, e^*_t)$.

With this assumption, the correlation between actions in different periods comes from the evolution of the underlying equilibria. Thus, actions in different periods are independent conditional on the 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:

$$Pr(a_{t+1}, a_t) = \sum_{e^*_{t+1}} Pr(a_{t+1}, e^*_{t+1}, a_t) = Pr(a_{t+1}|e^*_{t+1})Pr(a_t|e^*_{t+1})Pr(e^*_{t+1}) \tag{2.7}$$

where $Pr(a_l|e^*_{t+1})$ represents the probability of the players choosing action a_l 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,

$$\Pr(a_{t+2}, a_{t+1}, a_t) = \sum_{e^*_{t+1}} \Pr(a_{t+2}|e^*_{t+1}) \Pr(a_{t+1}|e^*_{t+1}) \Pr(a_t|e^*_{t+1}) \Pr(e^*_{t+1}) \tag{2.8}$$

Sum over a_{t+1} leading to:

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

Based on the above two equations connecting the observed and unknowns with the identified of number of equilibria, first I can partition the actions into Q alternatives. Then CCPs under each equilibrium can be identified as eigenvalues of the observed matrices consisting of joint distributions (see Appendix for detail). With tackling the multiple equilibria problem, the payoff functions can be nonparametrically identified with exclusion restrictions, exactly the same as the cross-sectional case. As a result, I summarize the identification results with panel data in the following theorem:

Theorem 2.11. ***(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.*

2.5 Semi-parametric Estimation

Estimation is implemented in two steps. First I present estimation of all equilibrium-relevant aspects such as the number of equilibria and the equilibrium CCPs. Then I explained the estimation of payoff structural parameters through minimum distance estimation.

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[6]. 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

[6]See Newey (1990), Ai and Chen (2003) and Newey (1994) for how to use sieve series expansions to estimate.

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 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:

$$\hat{Pr}(a_1, ..., a_n|x) = \frac{\frac{1}{M}\sum_m I(a_1^m = a_1, ..., a_n^m = a_n, x^m = x)}{\frac{1}{M}\sum_m I(x^m = x)}$$

With estimation of the joint distributions, matrix F_{a_1,a_2} is obtained since it essentially collects all joint distributions of actions of player 1 and 2.

The rank of matrix F_{a_1,a_2} can be used to as an estimator for the number of equilibria as I show in the identification. To estimate the rank of a matrix, this paper follows the procedure developed in Kleibergen and Paap (2006) to test the null hypothesis that the rank of the matrix is equal to a given integer. There are a lot of ongoing research on estimating the rank of a matrix[7].

To test whether the rank of a $h \times l$ matrix A equals to r, one can use the number of non-zero singular values. For a generic matrix A, its singular value decomposition can

[7]See the 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

be expressed as:

$$A = USV' = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix} \begin{pmatrix} V'_{11} & V'_{12} \\ V'_{21} & V'_{22} \end{pmatrix}$$

where U is an $h \times h$ orthogonal matrix, V is a $l \times l$ orthogonal matrix, and S is an $h \times l$ matrix that contains the singular values of A in decreasing order on its main diagonal and is equal to zero elsewhere. In the partition of U, S and V, U_{11}, S_1 and V_{11} are $h \times h$, and the dimensions of other submatrices are defined accordingly. With this decomposition and partition, the null hypothesis $H_0 : rank(A) = r$ is equivalent to $H_0 : S_2 = 0$ because a matrix's rank is defined as the number of non-zeros singular values.

The statistic proposed in Kleibergen and Paap (2006) utilizes an orthogonal transformation of S_2 as follows. $\Lambda_r = A'_r A B'_r$, where $A'_r = (U_{22}U'_{22})^{1/2}(U'_{22})^{-1}[U'_1 2U'_{22}]$ and $B_r = (U_{22}U'_{22})^{1/2}(U'_{22})^{-1}[U'_{12}U'_{22}]$. the null hypothesis $H_0 : rank(A) = r$ is equivalent to $H_0 : \Lambda_r = 0$. Let $\hat{A}$ be an estimator of the matrix A with sample size N, to derive a nice asymptotic distribution for the test, the estimator of the column vectorization of matrix A is asymptotical normal, which is satisfied since in our case each element in the matrix is a frequency estimator. From theorem 1 in Kleibergen and Paap (2006), $\hat{\lambda}_r$ is asymptotically normal distributed. To estimate the rank of the population rank of A, a sequential test needed to implement such as test $H_0 : rank(A) = r$ against $H_1 : rank(A) > r$ starting from $r = 1$, then $r = 2, ..., min\{s, t\}$. The first value for r that leads to a nonrejection of H_0 generates the estimate for 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.

When the number of equilibria is known, the estimation of the CCPs under different equilibria exactly follows the identification procedure. For instance, with information from three players, CCPs of one player can be estimated as the eigenvector from the matrix decomposition and through matrix manipulation.

2.5.1 Parametric Estimation of the Payoff Function

With CCPs under each equilibrium estimated, payoff functions can be estimated nonparametrically with exclusion restrictions by following the identification procedure. However, in practical, nonparametric estimation poses a very high demanding on the data. I therefor parameterize the payoff function and estimate the structural parameters instead. Denote the parameterized payoff functions as $\pi_i(a_i, a_{-i}, s) = \pi_i(a_i, a_{-i}, x; \theta)$.

Pioneered by Hotz and Miller (1993)[8][9], 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 the fixed point. It is well known that looking for a

[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.

[9]See also Su (2012) for a novel constrained optimization method

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 $\tilde{\sigma}(a|x, e^*)$. The equilibrium condition is represented by a general mapping denoted as $h(\sigma, \theta) = \sigma(a|x, \theta) - \Gamma(\sigma(a|x, \theta); \theta) = 0$, which holds for every x. The least squares estimator estimates the parameters of interest by forcing the constraints:

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

satisfied approximately for every x and every equilibrium. With the number of equations greater than the number of parameters, a weight is assigned to individual equations for minimization. Denote $\tilde{\sigma}_M$ as the vector of collecting all $\hat{\sigma}(a|x, e^*)$ and $\Gamma(\hat{\sigma}; \theta)$ as another vector collects all $\Gamma(\hat{\sigma}(a|x, e^*); \theta)$. Let W_M be a symmetric positive definite matrix with dimension of $((K+1)^n \cdot \sum Q_x) \times ((K+1)^n \cdot \sum Q_x)$ 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

$$\hat{\theta}(W_M) = argmin_{\theta} \quad [\hat{\sigma} - \Gamma(\hat{\sigma};\theta)]' \, W_M \, [\hat{\sigma} - \Gamma(\hat{\sigma};\theta)]$$

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.

2.6 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 x. Assume the market attribute is discrete. Thus, market characteristics x 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:

$$\begin{aligned}\pi(a_i = 1, a_{-i}; x) &= \beta x + \delta \frac{\sharp(a_{-i} = 1)}{n-1} + \epsilon_{i1} \\ \pi(a_i = 0, a_{-i}; x) &= \delta \frac{\sharp(a_{-i} = 0)}{n-1} + \epsilon_{i0}\end{aligned}$$

where ϵ_{i1} and ϵ_{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 $\beta = 0.04$, $\delta = 2.5$ and only considering symmetric equilibria, all the equilibria for different market types $x = 1, 2, 3, 4$ are presented in figure A.3. Regarding market types $x = 1, 2, 3$, there are three equilibria, among which the middle one is unstable, while there is a unique equilibrium for market type $x = 4$.

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 $500, 800, 1000, 1500, 2000, 4000, 6000, 8000, 10000$. The reported results are based on 1000 simulated samples from mixtures with two equilibria. From figure A.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 Estimation of CCPs bases on information of the number of equilibria. If the number of equilibria is unknown and needs to be estimated, the estimation then serves as a model selection procedure. Even though the selection can constructed so that it is consistent in large samples, the estimation of CCPs is conducted in the same data within which the selection is implemented. To control for the post-section inference is out of the scope of this paper. Also this Monte Carlo evidence is to show the performance of the CCPs' estimates through eigenvalue-eigenvector decomposition. Thus, 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 A.1 and table A.3). CCPs are estimated with high accuracy. 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 A.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 x, 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 not 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 θ using CCPs of any equilibrium set should yield the same estimates.

Suppose the econometrician observes x 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 ($\{\Pr(a_i = 1|x = 1, 2, 3, \tau = 1), \Pr(a_i = 1|x = 1, \tau =$

$2), \Pr(a_i = 1|x = 4)\})$ and another set of CCPs $(\{\Pr(a_i = 1|x = 1, 2, 3, \tau = 2), \Pr(a_i = 1|x = 2, \tau = 1), \Pr(a_i = 1|x = 4)\})$ separately. Using CCPs from both equilibria of some values of x is for identification purposes. When players are identical, the parameters are not identified without multiplicity. The testing fails to reject the null hypothesis that only 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 $x = 1, 4$, among which there are two equilibria selected in market $x = 1$. 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 δ, and test whether this δ is the same or not. The testing rejects the null hypothesis that the structural primitives estimated are the same.

2.7 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.

2.7.1 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 A.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 A.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 (2009), 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 A.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 Sweeting (2009).

2.7.2 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 (2009), station i's payoff for placing a commercial in time block $t \in \{0, 1\}$ is defined as follows:

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

where p_{-it} is the proportion of stations in the same market choosing timing t. α_t allows different average profit for stations when they play their commercials in different timing t, and δ captures the coordination incentives. Stations receive the idiosyncratic private profit shocks before they make their timing decisions. The ϵs represent the fact that a station tends to play commercials at different times every day are represented by ϵ_{it}. This variation is because the length of other programming, such as songs or travel news, can vary and be unpredictable. A station would not want to annoy its listeners by cutting short other programs to play commercials at precise times.

As usual, I assume ϵ_{it} to be i.i.d with extreme value distributions across actions, players and markets. α_0 is normalized to be zero for identification purposes. Denote α_1 as α 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 δ is identified, and α is identified.

2.7.3 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 (2009).

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 (2009), 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 α 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 α 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 αs 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.

2.8 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 equilibria and the equilibrium selection mechanism are identified in the static game setting. 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. 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. Thus, it is important to provide identification while consider both unobserved heterogeneity and multiple equilibria together. Another direction for future research is to provide identification and estimation for dynamic games while consider multiple equilibria and/or market unobserved heterogeneity. Dynamic games are increasingly used to investigate the strategic interaction among firms over time while identification is not clear when multiplicity of equilibria and/or unobserved market-level heterogeneity exist. Without the identification, one would not be confident of obtaining the right estimates for the underlying structural.

Chapter 3

Identification of Incomplete Information Games with Multiple Equilibria and Unobserved Heterogeneity

3.1 Introduction

Games are widely used to investigate strategic interactions between players in industrial organization. Unlike single-agent discrete choice models, games generally admit multiple equilibria. Allowing for the presence of multiple equilibria complicates the identification and estimation for games, while ignoring it may result in mis-specification. However, if the data is generated by the same equilibrium, as assumed in the conventional assumption, the two-step estimators pioneered by Hotz and Miller (1993) enables consistent estimation the structural parameters. The two-step estimators work without requiring researchers knowing which equilibrium being employed. Moreover, there is no need to solve for all the equilibria. However, there is hardly empirical evidence on this conventional assumption.

Meanwhile, another potential issue existing in estimation of games is that econometrician might not get access to all the covariates that affect players' payoffs, meaning that there is some factors either in the market level or individual player level referred as unobserved heterogeneity. Again an easy way to get around the difficulty created by missing information is to ignore it by pretending that econometricians observe all relevant information. This treatment is unrealistic for most applications in empirical IO and also problematic to explain micro data. Not accounting for unobserved heterogeneity can generate significant biases in parameter estimates, [1], thus misleading when economists explain the strategic interactions between firms. Consequently, addressing unobserved heterogeneity is also important in empirical estimation since it is widely incorporated in empirical estimations. Meanwhile, accounting for unobserved heterogeneity, and therefore dynamic selection, is also important in dynamic discrete choice models in labor economics.

Without imposing assumptions on the equilibrium selection and allowing for unobserved market level heterogeneity, I provide identification for finite action static games with

[1] For instance, in the empirical application in Aguirregabiria and Mira (2007), the estimation without unobserved market heterogeneity implies estimates of strategic interaction between firms (i.e., competition effects) that are close to zero or even have the opposite sign to the one expected under competition. While including unobserved heterogeneity in the models results in estimates that show significant and strong competition effects

incomplete information [2]. From Chapter 2, with traditional assumption that the private shocks are independent, the presence of multiple equilibria yields a finite mixture structure over the observed joint distribution of actions and equilibrium choice probabilities. On the other hand, the existence of unobserved market level factor also creates a similar finite mixture structure if the factor is finite. Consequently, allowing for both latent factors creates the similar finite mixture feature as represented in Chapter 2. Then the cardinality of the combination of multiple equilibria and unobserved market types can be identified as the rank of a matrix constructed by joint distribution of actions. In addition, conditional choice probabilities (CCPs) can be identified as eigenvectors of the matrix through decomposition. The biggest problem for the identification is to distinguish between multiple equilibria and the unobserved market types. Without ordering of the CCPs, the structural parameters can only be set identified instead of point identified.

Incorporating both multiple equilibria and unobserved heterogeneity is important for estimation in empirical studies. Existing literature usually consider one possibility while assume the other possibility away. For instance, De Paula and Tang (2012) infer the sign of strategic interaction term via assuming only multiplicity of equilibria is present. In contrast, lots of literature take into account unobserved heterogeneity in estimation but make assumption that the data is generated by the same equilibrium, e.g., Aguirregabiria and Mira (2007). As far as I know, besides this paper, there is only one paper by Aguirregabiria and Mira (2013) that considers both latent factors together. This paper provides conditions that makes the argument in Aguirregabiria and Mira (2013) complete.

I organize the rest of the paper as the follows. I begin with outline of the static

[2] Examples of such games include radio stations commercial airing decision in Sweeting (2009)

game framework in section 2. Set identification results are provided in section 3. Section 4 provides a Monte Carlo evidence. Then section 5 concludes. The Appendix contains the proofs, the figures and the tables.

3.2 Game Setup

Consider a static simultaneous move game that involves N players. Players obtain action specific 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, $i \in \{1, ..., N\}$, chooses an action a_i out of a finite set $\mathscr{A} = \{0, 1, ..., K\}$. Let a_{-i} denote player i's rivals' actions and $x \in \mathscr{X}$ denote public observable state variable. The $K+1$ action specific profit shocks are denoted as $\epsilon_i(a_i)$, and their density distributions are denoted as $f(\epsilon_i)$[3]. The payoff for player i from choosing action a_i is assumed to be additive separable as below:

$$U_i(a_i, a_{-i}, x, \epsilon_i) = \pi_i(a_i, a_{-i}, x) + \epsilon_i(a_i)$$

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

Instead of defining the equilibrium using players' decision rules, I defined the equilibrium using the probability that each player choosing each possible action, i.e. $\sigma_i(a_i|x)$ denotes the probability that player i chooses action a_i conditional on observing x. Since player i does not observe her rivals' payoff shocks, she has to form belief over the distribution

[3] Similar setups are studied in Seim (2006) and Aradillas-Lopez (2010).

of actions that her rivals are going to choose. Meaning, player i needs to best response to her rivals' action distribution instead of a specific action. With the following independence assumption, player i's belief does not depend on her own private information.

Assumption 3.1. *(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.*

The assumption of conditional independence among 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. Consequently, player i expects her rivals' action distribution to be $\sigma_{-i}(a_{-i}|s)$. Then I can represent player i's expected utility from choosing action a_i as the following:

$$u_i(a_i, x, \epsilon_i) = \sum_{a_{-i}} \pi_i(a_i, a_{-i}, x)\sigma_{-i}(a_{-i}|x) + \epsilon_i(a_i) \equiv \Pi_i(a_i, x) + \epsilon_i(a_i)$$

The Bayesian Nash Equilibrium is stated in the following:

Definition 3.2. ***(BNE)*** *For a fixed state s, the Bayesian Nash Equilibrium (BNE) is a collection of probabilities $\sigma_i^*(a_i = k|x)$ for $i = 1, ..., N$ and $k = 0, ..., K$ such that for all i and k, the following equation satisfied:*

$$\sigma_i(a_i = k|x) = \Pr\left(\Pi_i(a_i = k, x) + \epsilon_i(a_i = k) > \Pi_i(a_i = j, x) + \epsilon_i(a_i = j), \forall j\right)$$

Following Hotz and Miller (1993), the equilibrium condition implies a one-to-one mapping between the CCPs $(\{\sigma_1^*(a_1|x), ..., \sigma_n^*(a_n|x)\})$and the difference of expected choice utilities $(\{\Pi_1(a_1, x) - \Pi_1(a_1 = 0, x), ..., \Pi_n(a_n, x) - \Pi_n(a_n = 0, x)\})$

Assumption 1 guarantees that the mapping has at least one fixed point by Brouwer's fixed point theorem. As a result, if CCPs are obtained, the differences of the expected choice values are as well. If there is a unique equilibrium, CCPs computed from the collected data can be used to approximate the CCPs predicted by theory. Then 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 equilibria. 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 players favor one equilibrium over the other. 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.

3.3 Set Identification

This section investigates the identification of the static game with incomplete information described above. The importance of the identification is to incorporate unobserved market-level heterogeneity and allow for multiple equilibria. First I provides conditions under which the cardinality of the overall latent variable, combining information from both unobserved heterogeneity and multiple equilibria can be uniquely recovered conditional on market observables. Next, the conditional choice probabilities associated with each level of the overall latent variable can be identified up to ordering. Thus point identification requires extra information to disentangle the two types of unobserved factors. Without imposing very restrictive assumption, a set identification is proposed for the conditional choice probabilities and hence the structural parameters.

Before going into details of the identification, I describe the data structure first. Suppose the econometrician observes actions of all players in cross-sectional markets m where $m = 1, ..., M$ with characteristics $x_m \in X \equiv \{x^1, ..., x^d\}$. The state variable space X is assumed to be discrete and has a finite support d. The unobserved market-types are denoted as η, which is also assumed to be finite and have a finite support. Assuming that the number of equilibria is also finite, index the equilibrium as $e^* \in \Omega_{x,\eta} \equiv \{1, ..., Q_{x,\eta}\}$, where $Q_{x,\eta}$ is the number of equilibria being employed in market with characteristics $\{x, \eta\}$[4]. Note

[4]If the number of equilibria is infinite, it is impossible to do the identification using limited information.

that the cardinality of the equilibrium set varies with market observed characteristics x and the unobserved market type η. Since the identification of equilibrium-specific components is established conditional on the market characteristics x_m, I suppress it for ease of notation. I will reintroduce market characteristics when I move to the identification of payoff functions.

The joint distribution of individual players' actions can be expressed as the following:

$$\begin{aligned}
\Pr(a_1, ..., a_n) &= \sum_{\eta} \Pr(a_1, ..., a_n|\eta) \Pr(\eta) && (3.1)\\
&= \sum_{\eta} \sum_{e^*} \Pr(a_1, ..., a_n|e^*, \eta) \Pr(e^*|\eta) \Pr(\eta)\\
&= \sum_{\eta} \sum_{e^*} \Pr(a_1|e^*, \eta)\times, ..., \times \Pr(a_n|e^*, \eta) \Pr(e^*, \eta)\\
&\equiv \sum_{\tau} \Pr(a_1|\tau)\times, ..., \times \Pr(a_n|\tau) \Pr(\tau) && (3.2)
\end{aligned}$$

The first and second equality are due to the law of total probability, where $\tau \equiv \{e^*, \eta\}$ is the newly created latent variable that aggregates information from both the unobserved market types and the equilibria associated with each market type. Denote the cardinality of τ as Q_x which can be computed as $Q_x = \sum_{\eta} Q_{x,\eta}$. CCP $\Pr(a_i|\tau)$ represents the probability player i choosing action a_i in market type η and equilibrium e^*. The cardinality of the new latent variable τ can be identified following the same intuition in the first chapter of this dissertation.

3.3.1 Point Identification of the Cardinality

Similar to the scenario without unobserved market types in Chapter 2, the cardinality of the combined latent factor can be obtained using correlation between two players

Also by (Harsanyi (1973)), we know that the game that has infinite number of equilibria has zero measure through their joint distribution of actions. With two players, the joint distribution of actions becomes:

$$\Pr(a_1, a_2) = \sum_{\tau} \Pr(a_1|\tau) \Pr(a_2|\tau) \Pr(\tau) \tag{3.3}$$

Matrixes are used extensively in this paper to make use of all possible information together at the same time. I rewrite this equation into a matrix form:

$$F_{a_1,a_2} = A_{a_1|\tau} D_\tau A_{a_2|\tau}^T \tag{3.4}$$

where

$$\begin{aligned} F_{a_1,a_2} &\equiv [\Pr(a_1 = k, a_2 = j)]_{k,j}\,, \\ A_{a_i|\tau} &\equiv [\Pr(a_i = k|\tau = q)]_{k,q} \\ D_\tau &\equiv \operatorname{diag}\Big[\Pr(\tau = 1) \ \ldots \ \Pr(\tau = Q)\Big]. \end{aligned}$$

Those matrices stack the distributions with all possible values that a_1, a_2 and τ can take. In particular, matrix F_{a_1,a_2} consists of the whole joint distributions of a_1 and a_2, which can be estimated from data. D_τ is a diagonal matrix with the marginal distribution of the new latent variable τ as the diagonal elements, while matrix $A_{a_i|\tau}$ collects all the CCPs associated with equilibrium τ.

The dimensions of the three matrices defined above F_{a_1,a_2}, $A_{a_i|\tau}$ and D_τ 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 3.3. ***(Identification of the Cardinality)*** *The rank of the observed matrix* F_{a_1,a_2} *serves as the lower bound for the number of equilibria, i.e.,* $Q \geq Rank(F_{a_1,a_2})$*. Furthermore, the number of equilibria is identified, particularly,* $Q = Rank(F_{a_1,a_2})$ *if the following conditions are satisfied:*

(1) $K+1>Q$*; (2) both matrices* $A_{a_1|\tau}$ *and* $A_{a_2|\tau}$ *are full rank; (3) all* $\Pr(\tau)$ *are positive.*

Proof The proof is the same as lemma 2.3 □

The first condition requires that the number of possible actions 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. However, here I only use data from two players. If there are more players, grouping all players into two groups expands the action space that a representative player can choose from. Consequently, it increases identification power.

The full rank condition implies that no equilibrium or market type is redundant. The third condition indicates that only those equilibria that are active in the data 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 does not affect identification and estimation consistency.

The cardinality identified as described above indicates the feature of the underlying latent factor to some extent. Specifically, if the cardinality varies with the market characteristics, then we are comfortable to conclude that multiple equilibria exist in the data. A typical assumption imposed on the market-level unobserved heterogeneity is that the support is fixed across different market observables due to the fact that it is exogeneous.

In contrast, the number of equilibria employed varies with market characteristics because it is endogeneous. However, a universal cardinality of the latent variable does not necessary preclude the existence of multiple equilibria. For instance, it could be the fact that each market employs the same number of equilibria in the data.

3.3.2 Identification of CCPs

Provided that the cardinality of the unobserved variable is identified, I show below how to identify CCPs up to ordering for each individual player. To utilize the invertible of matrices, I partition the action space so that it has the same dimension as the cardinality of the overall latent variable τ. With a little bit abuse of notation, I still use a_1 and a_2 to represent the action variable after the partition. One key criteria for the partition is that the corresponding tailored matrix $A_{a_i|e^*}$ is full rank for both $i = 1, 2$. The partition is feasible because the original matrix is full column rank.

Identification for individual players' equilibrium strategies requires extra information. I denote this additional player as player 3. Then the following equation links observed joint distribution of actions with those unknowns:

$$\Pr(a_1, a_2, a_3) \quad = \quad \sum_{\tau} \Pr(a_1|\tau) \Pr(a_2|\tau) \Pr(a_3|\tau) Pr(\tau)$$

To use the matrix algebra for identification, I fix $a_3 = k$ and use all possible variation from player 1 and 2's actions. Matrices definition is the same as I defined above with the only difference is that a_3 is fixed. Using the above matrix representations, we have the following two equations:

$$F_{a_1,a_2} = A_{a_1|\tau} D_\tau A_{a_2|\tau}^T \tag{3.5}$$

$$F_{a_1,a_2,a_3=k} = A_{a_1|\tau} D_{a_3=k|\tau} D_\tau A_{a_2|\tau}^T \tag{3.6}$$

Given that $A_{a_1|\tau}$ and $A_{a_2|\tau}$ have full rank, post-multiplying F_{a_1,a_2}^{-1} on both sides of equation 3.6 leads to the following main equation, which is essential for the identification.

$$F_{a_1,a_2,a_3=k} F_{a_1,a_2}^{-1} = A_{a_1|\tau} D_{a_3=k|\tau} A_{a_1|\tau}^{-1} \tag{3.7}$$

The right-hand side of the equation above represents an eigenvalue-eigenvector decomposition of the matrix on the left-hand side, with matrix $D_{a_3=k|\tau}$ being the diagonal matrix consisted of eigenvalues and matrix $A_{a_1|\tau}$ being the eigenvector matrix. Given that the left-hand side matrix in the equation can be estimated from the observed data, this equation can be used to identify both $D_{a_3=k|\tau}$ and $A_{a_1|\tau}$ simultaneously (see Hu (2008)). However, the matrix decomposition only can identify the eigenvalue matrix up to ordering and the eigenvector matrix up to normalization. Normalization can be obtained through the fact that the sum of each column eigenvector is one because it is consisted of a whole probability distribution (The identification is the same as provided in Chapter 2 in this dissertation).

Lemma 3.4. ***(Identification of the CCPs Up to Ordering)*** *With conditions in lemma 3.3 satisfied, the CCPs are identified up to ordering.*

Proof The proof is the same as lemma 2.6. □

3.3.3 Set Identification of Payoffs

With the equilibrium CCPs can only be identified up to ordering, I proceed to discuss how to utilize this information to recover the payoffs of individual players. If the equilibrium CCPs can be uniquely identified, so do the payoffs. Thus, to uniquely recover the payoffs, ordering of the equilibrium CCPs are very important. In particular, two things needed to be done for point identification of the payoffs since the CCPs are identified conditional on market observables x. First one has to distinguish which values of τ belongs to the same latent market-type. To be more specific, one has to divide the Q_x sets of CCPs into c groups, and each group contains all the equilibria employed. This generates c group conditional on market characteristics x. Secondly, one has to order the unobserved market types because they play an important role in payoffs.

The difficulty for point identification lies in how to disentangle the equilibrium and unobserved market types so that ordering of the latent variable τ is obtained. In discrete choice models with individual unobserved heterogeneity, usually a monotonicity assumption is imposed to order the latent variable, e.g., Hu, McAdams, and Shum (2013). The key difference between discrete choice model and games are the presence of multiple equilibria. Unfortunately, a similar monotonicity assumption is not feasible in the game setting even if we assume away multiplicity of equilibria by imposing assumption that the data is generated by the same equilibrium. With this assumption, the first difficulty mentioned above vanishes. However, Assuming that multiple equilibria do not exist in the data does not necessary rule out multiple equilibria exiting in the theoretic model. Moreover, this assumption does not specify which equilibrium is employed in the data. Therefore, monotonicity of CCPs is highly possible to be violated.

In figure B.1, I characterize the relationship between CCPs and both market observed and unobserved characteristics. To illustrate the idea but keep things simple, I assume there are two different latent market-types. In markets with observed characteristics x^1 and x^2, a unique equilibrium exists in market-type 1 while three equilibria in market-type 2. In market-type 2, the same equilibrium is employed in market with characteristics x^1 and x^2 respectively. However, note that when market characteristics differ, it indicates the games are different from each other. Thus, it is possible that in the data, the low equilibrium is employed in market with characteristics x^1 while the middle equilibrium is employed in market with characteristics x^2. Therefore, we have:

$$\Pr(a_i = k|x = x^1, \eta = 1) < \Pr(a_i = k|x = x^1, \eta = 2)$$

$$\Pr(a_i = k|x = x^2, \eta = 1) > \Pr(a_i = k|x = x^2, \eta = 2)$$

A straightforward fix to the above non-monotonicity could be an extra assumption on the equilibrium selection, in which both markets with characteristics x^1 and x^2 always select the same type of equilibrium. For instance, in the above example, either both low, middle or high equilibrium are employed at the same time. Unfortunately, this assumption is very restrictive because the number of equilibria varies with the market characteristics (see figure B.2.

Ordering using monotonicity conditions will be more difficult when multiple equilibria are allowed. This is because when we compare the conditional choice probabilities associated with two different values of τ, it is difficult to tell whether they are associated with different market types or just associated with two equilibria with the same market type. Moreover, the index of equilibrium has no economic meanings. In particular, equilibrium 1 in one market has nothing to do with equilibrium 1 in another market. Without additional assumptions, the decomposition provides us a whole set of choice probabilities without ordering, which will be a permutation of the underlying η for a given x.

The unidentified ordering prevents us from uniquely recovering the payoffs. Without information of which CCPs belongs to the same market type but different equilibria, it is impossible to recover how individual player's payoff differs in different market-types. The variation in probability of players choosing different actions in different market-types reveals information of how payoff differs in those different market-types, thus providing the identification power. In particular, we need to know that which two values of η are different equilibria associated with the same market type, and which two values of η associated with different market types so we can identify how payoff varies with different market types. Without imposing assumptions regarding the equilibrium selection and some monotonicity assumption, distinguish of multiple equilibria and unobserved types are infeasible. Thus point identification is also infeasible.

Without making further assumption, this paper pursues set identification instead of point identification, but I provide examples and conditions that point identification can be obtained. Similar in discrete choice models, it is impossible to separately identify all the payoff functions, only their differences can be recovered. Normalization and exclusion restrictions are necessary, and I state them as below:

Assumption 3.5. ***(Normalization)*** *For all i and all a_{-i} and s, $\pi_i(a_i = 0, a_{-i}, x, \eta) = 0$.*

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.

Assumption 3.6. ***(Exclusion Restriction)*** *For each player i, the state variable can be* *partitioned into two parts denoted as x_i, x_{-i}, so that only x_i enters player $i's$ payoff, i.e. $\pi_i(a_i = k, a_{-i}, x) \equiv \pi_i(a_i = k, a_{-i}, x_i)$.*

With the normalization and exclusion restriction assumptions, if private shocks follow extreme value distribution, the equilibrium condition becomes

$$log\sigma_i(a_i = k|x, \eta, e^*) - log\sigma_i(a_i = 1|x, \eta, e^*) = \sum_{a_{-i}} \pi_i(a_i = k, a_{-i}, x_i, \eta)\sigma_{-i}(a_{-i}|x, \eta, e^*)$$

The extreme value distribution is not necessary for the identification, but for illustration purpose. From above equation, if CCPs are identified, payoff functions can be non-parametrically identified. The exclusion restriction x_{-i} shifts the choice probability while keeping the payoff fixed, variation of the exclusion restriction identifies the payoff functions. (see Bajari, Hong, Krainer, and Nekipelov (2010b) and Bajari, Hahn, Hong, and Ridder (2011)).

However, as I stated above, the CCPs are identified up to ordering. Equivalently, the two unobserved types cannot be distinguished from the first step identification. It is hardly to find conditions to tell whether two sets of CCPs are associated with different equilibria but the same latent market types. However, one of the ordering is the correct one, which is equivalent to the fact that the set consisted of all possible orderings must contain the true ordering. Consequently, the set consisted of payoffs rationalized by CCPs of any possible ordering must also include the true payoffs. Express the first step CCPs as $\{\Pr(a_i = k|x, \tau), k = 0, ...K, i = 1, ..., n, \tau = 1, ..., Q_x, x \in \{x^1, ..., x^d\}\}$

Theorem 3.7. ***(Set Identification of the Payoffs)*** *With the cardinality of latent factors being identified and CCPs are identified up to ordering, the identified set for the payoff* *functions can be characterized as:*

$$\pi_i(a_i, a_{-i}, x, \eta) \in \Psi \equiv \{\pi_i(a_i, a_{-i}, x, \tilde{\eta}),\ \textit{that } \tilde{\eta} \textit{ is a possible ordering of the original } \tau\}$$

It is computationally light and easy to implement when conduct estimation following the identification set. It does not require to solve for the equilibria even once. Also, the total number of possible ordering is a finite number, generating an identified set with a finite support. The number of possible ordering can be computed in two steps. First, there are $C(Q, c)$ possible ways to divide Q number of values into c groups. Secondly, order those groups across different market characteristics results in a total number of ordering to be $\Pi_x C(Q, c)$

Illustration of the set identification. I provide a simple example to illustrate the essential of the set identification for structural parameters θ. Assume that both observables x and unobservables η have a support of 2, denoted as x^1, x^2 and $\eta = 1$, $\eta = 2$ respectively. In the data generating process, one equilibrium is employed in market $\eta = 1$ and two equilibria are employed in market $\eta = 2$. To summarize, for $x^1 : \eta = 1, (\eta = 2, e^* = 1), (\eta = 2, e^* = 2)$ and for $x^2 : \eta = 1, (\eta = 2, e^* = 1), (\eta = 2, e^* = 2)$. Therefore, when we identify the cardinality to be 3, we know that there are a unique equilibrium associated with one η and two equilibria associated with another η, but we do not know exactly that markets with which η employs the unique equilibrium.

The possible ordering of the three sets of CCPs (P^1, P^2, P^3) are characterized as follows:

	x^1		x^2	
θ	$\eta = 1$	$\eta = 2$	$\eta = 1$	$\eta = 2$
$\hat{\theta}^1$	P^1	(P^2, P^3)	P^1	(P^2, P^3)
$\hat{\theta}^2$	P^1	(P^2, P^3)	P^2	(P^1, P^3)
$\hat{\theta}^3$	P^1	(P^2, P^3)	P^3	(P^1, P^2)
$\hat{\theta}^4$	P^2	(P^1, P^3)	P^1	(P^2, P^3)
$\hat{\theta}^5$	P^2	(P^1, P^3)	P^2	(P^1, P^3)
$\hat{\theta}^6$	P^2	(P^1, P^3)	P^3	(P^1, P^2)
$\hat{\theta}^7$	P^3	(P^1, P^2)	P^1	(P^2, P^3)
$\hat{\theta}^8$	P^3	(P^1, P^2)	P^2	(P^1, P^3)
$\hat{\theta}^9$	P^3	(P^1, P^2)	P^3	(P^1, P^2)

So the identified set can be expressed as $\{\theta^1, ..., \theta^9\}$, and the a simple estimator can be $\{\hat{\theta}^1, ..., \hat{\theta}^9\}$. In what follows, I try to provide some conditions that point identification is obtained.

3.3.4 Point identification

As I discussed above, the undecided ordering of the unobserved market-type and the associated multiple equilibria precludes point identification of payoff functions. If there is no unobserved market level heterogeneity, ordering of equilibria does not matter because there is no economic meaning for the index of equilibria, and different sets of choice probabilities associated with different equilibria map to the same payoff. Point identification therefore can be obtained.

When unobserved market heterogeneity is incorporated into the model, can one obtain point identification if the data is free of multiplicity issue? For example, as the existing literature assumes that the data is generated by the same equilibria. From the graph illustrated above, unfortunately, we need assumptions more than that. The essential difference between market-level unobserved heterogeneity and multiple equilibria is that payoff varies with the values of the market-type while payoffs should be the same for different equilibria. Thus, in order to identify the payoff with the assumption that the data is generated by the same equilibria, we still need to order the market-level unobserved heterogeneity. Single agent discrete choice literature usually makes monotonicity assumption to order unobserved factor, i.e. the probability that an individual selects a certain actions increases or decreases with the value of the unobserved factor conditional on both market observed and unobserved covariates. However, in the game setup, it is questionable whether we can make similar monotonicity assumption. The assumption only rules out that when there are several equilibria, the same equilibrium is always employed. A stronger assumption is needed, and I state it in the following:

Assumption 3.8. *(i.) There exist some values of x such that a unique equilibrium is guaranteed for all η. (ii)* $\Pr(a_i|x, \eta)$ *is monotone with respect to η for all those xs.(iii.) For those xs that the game admit multiple equilibria for at least one of η, the equilibrium selection is not degenerated for at least one of the η.*

With this assumption, first of all, we can identify the cardinality of the unobserved market type L through $L = \min_x\{Q_x\}$. Furthermore, using information from those x that guarantees a unique equilibrium, the monotonicity assumption in *ii* can be used to order η. Thus, we can use those identified choice probabilities to recover the payoff function for those x. We can only identify payoffs for those x with a unique equilibrium nonparametrically. However, if we are willing to impose some parametrical assumption on the payoff structure, we probably can recover the structure parameters using only partial information. With the structural parameters being identified, we can solve all the equilibria for those x admitting multiple equilibria with different levels of market-types and then compare the choice probabilities to recover the equilibrium selection mechanism.

Corollary 3.9. ***(Point identification)*** *With assumption 3.3, 3.8, 3.5 and 3.6, the cardinality of unobserved market types, CCPs associated with different observed market types and the payoff functions can be point identified.*

It is obvious that the conditions for point identification is very restrictive, and it is very difficult to be satisfied in empirical studies. On the other hand, set identification sometimes is not satisfying and also hard to do inference. Moreover, given that the set identified above consists of finite number of vectors which itself is not convex. The existing literature requires the identified set to be convex for statistical inference. However, there is one way at least we can check whether there are only multiple equilibria or both multiplicity of equilibria and unobserved heterogeneity exist in the data.

As stated in Chapter 2, when only multiple equilibria are present, the payoffs are nonparametrically identified with exclusion restrictions. Moreover, existence of multiplicity of equilibria indicates that different equilibrium CCPs should map to the same payoffs since payoffs are equilibrium invariant. Consequently, a natural way to test is to set the null hypothesis as that only multiplicity of equilibria occur in the data. Under this null, the payoffs are nonparametrically identified and should be the same regardless which equilibrium CCPs are used to do the estimation. Consequently, testing whether payoffs are the same when different equilibrium CCPs are used is equivalent to the null that only multiple equilibria are present. Under the null, estimates of the structural parameters from minimum distance of the choice probabilities and its best response are asymptotically normal. Consequently, a statistic can be formed and its asymptotically chi-square distributed.

3.3.5 Alternative Set Identification

Assuming that the cardinality of the market unobserved heterogeneity is known denoted as J, this section studies set identification of static games with incomplete information allowing for multiplicity of equilibria. Besides difficulty of disentangling equilibria and unobserved heterogeneity, lack of variation also makes point identification infeasible. When there is not enough action options, only the lower bound of the cardinality can be identified. For instance, in a 2×2 entry game, using the methodology provided in this paper, one can only conclude that the cardinality of overall latent variable is greater than 1. Consequently, all components follows cannot be uniquely recovered. Moreover, using rank inequality to infer the cardinality imposes a full rank condition on the unobserved matrix, which is empirically not testable. With those concerns, a natural direction is to provides partial identification without imposing any of those assumptions.

Suppose the payoff functions are known up to a vector parameter θ, i.e. $\pi(a_i, a_{-i}, x, \eta) \equiv \pi(a_i, a_{-i}, x, \eta; \theta)$. The equilibrium set can be denoted as $\varepsilon(x, \eta, \theta)$, which maps a set of covariates and the parameters to a finite set of equilibrium strategy profiles $\varepsilon(x, \eta, \theta)$. The rank of the matrix constructed by the joint distribution of players' actions provides a lower bound to the overall cardinality which can be computed as $\sum_{\eta} \varepsilon(x, \eta, \theta)$. Every equilibrium profile implies a multinomial distribution over outcomes. If there is no unobserved heterogeneity, and $e^* \in \varepsilon(x, \eta, \theta)$ is the equilibrium being selected, the probability of observing outcome a if e^* is,

$$\tilde{\Psi}(a|x, \eta, \theta, e^*) \equiv \prod_i \prod_k \Pr(a_i = k|x, \eta, \theta, e^*)^{I(a_i)==k}$$

If both x and η are observed, and equilibrium is unique, $\tilde{\Psi}(a|x,\eta,\theta,e^*)$ could be compared directly to the observed data $\Pr(a|x,\eta)$. When there are multiple equilibria and no information of η, the observed outcome distribution is a mixture of equilibrium strategies according to a valid equilibrium selection mechanism $\lambda(x,\eta,\theta)$ and the marginal distribution of η λ^η, the probability of outcome a can be expressed as:

$$\tilde{\Psi}(a|x,\theta) \equiv \sum_{\eta} \lambda^\eta \sum_{e^*} \lambda(x,\eta,\theta) \prod_i \prod_k \Pr(a_i = k|x,\eta,\theta,e^*)^{I(a_i)==k} \tag{3.8}$$

Without imposing strong restrictions on the selection mechanism while allows missing information for point identification, I instead allows λ to be any valid mixture across equilibria and marginal distributions respectively, and derive an identified set implied by the model[5].

Theorem 3.10. *The identified set for the structural parameters θ for the static game can be represented as:*

$$\Theta_I = \left(\begin{array}{ll} \theta \in \Theta : \forall \quad a,x,\ s.t. & \Pr(a|x,\theta) = \tilde{\Psi}(a|x,\theta) \\ & \sum_\eta \lambda^\eta = 1 \quad and \quad \lambda^\eta \geq 0 \\ & \sum_{e^*} \lambda(x,\eta,e^*) = 1 \quad and \quad \lambda(x,\eta,e^*) \geq 0 \\ & \mid \varepsilon(x,\theta) \mid \geq \underline{Q}(x) \end{array} \right)$$

where $\tilde{\Psi}(a|x,\theta)$ is defined in equation 3.8. The last restriction uses information from the rank testing which provides a lower bound for the cardinality. 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 θ. Those θ that yields a number of equilibria smaller than the lower bound are excluded in the identified set. In the above 2×2 entry game example, if

[5]Similar identified set is provided in Grieco (2011) which allows flexible information structure.

the number of equilibria is 1, then point identification is obtained; if the number of equilibria is greater than 1, then those θ which provides a singleton equilibrium are deleted from the identified set.

Likelihood Representation of the Identified Set

Denote the true data generating process with $\{\theta^0, \lambda^0 = (\lambda^0_\eta, \lambda^0(x, \eta, e^*))\}$. The true model is complete in the sense that it includes a well defined model, a valid marginal distribution and a valid equilibrium selection mechanism, with which the model maps onto a unique point in the space of outcome distribution of a given x. The partial identification issue arises because there maybe multiple parameters (θ', λ') that generates the same conditional outcome distribution as (θ^0, λ^0).

The sample log-likelihood functions can be written as:

$$L_M(\theta, \lambda) = \frac{1}{M} \sum_m log(\tilde{\Psi}(a|x, \theta, \lambda)$$

the limit of the log-likelihood function, $L(\theta, \lambda) = E[log(\tilde{\Psi}(a|x, \theta, \lambda)]$, will be maximized at (θ^0, λ^0). However, without point identification, the maximizer is highly likely to be multiple, i.e. there may exist (θ', λ') such that $L(\theta^0, \lambda^0) = L(\theta', \lambda')$. As usually we are more interested in the payoff structural parameters θ^0, here I treat the λ as a nuisance parameter and focus on θ as the object of interest. The identified set can be represented as the set of maximizers of L.

$$\Theta_I \equiv argsup_{\theta \in \Theta} \sup_\lambda L(\theta, \lambda)$$

Therefore, the identified set can be formed through a pair of (θ, λ) that attains the maximal of the sample log-likelihood function.

3.4 Monte Carlo Simulation

This section presents some Monte Carlo evidence to illustrate the proposed identification methodology in the static game framework. Suppose n players decide to enter or stay out of markets with characteristics x. Assume the market attribute is discrete. Thus, market characteristics x can be regarded as the market observed type. Assume players are homogeneous, and the payoff functions of entry (1) or not (0) are parameterized as follows:

$$\begin{aligned}\pi(a_i = 1, a_{-i}; x, \eta) &= \beta x + \delta \frac{\sharp(a_{-i} = 1)}{n-1} + \eta + \epsilon_{i1} \\ \pi(a_i = 0, a_{-i}; x, \eta) &= \delta \frac{\sharp(a_{-i} = 0)}{n-1} + \epsilon_{i0}\end{aligned}$$

where ϵ_{i1} and ϵ_{i0} are private shocks, and η is the unobserved market heterogeneity. Assuming the private information is independent and identically follows extreme value distribution. Given this specific payoff function, the number of players does not affect individual player's strategy, i.e., only the fraction of players entering matters. Assume that ϵs are i.i.d across player and actions, and follows type 1 error distribution. Note that players are homogeneous. I only consider symmetric equilibria in this simulation. let $p(x, \eta)$ denote the probability an individual player chooses action 1 in the equilibrium conditional on market characteristics x, η.

$$log(p(x, \eta)) - log(1 - p(x, \eta)) = \beta x + \delta(2p(x, \eta) - 1) + \eta$$

For $\beta = 0.04$, $\delta = 2.5$, all the equilibria for different observed and unobserved market types are presented in figure B.3. The number of equilibria varies with both x and η continuously. η captures the unobserved factor that affects payoff of choosing action 1. A positive η also generates the same clustering effect from coordination by positive δ. Ignoring the existence of η might conclude the coordination effect is higher than it is true value, which is misleading to understand the feature of the markets.

Consider only symmetric equilibria, I solve for all the equilibria for different combination of x and η and presents how the equilibrium probability of choosing action 1 varies with x and η (see figure B.3). The number of equilibria varies non-monotonically with both x and η, it is therefore hard to infer the relation between the number of equilibria and market characteristics without knowing the underlying payoff parameters. Moreover, mixing with multiple equilibria, the equilibrium probability of choosing action 1 does not have the monotonicity along with η.

The Monte Carlo experiment consists of repetition of 500 with sample sizes of 800,1000,1500,3000,5000 for markets with different values of x and η. For each replication, I generate the equilibrium employment according to the equilibrium selection mechanism, and then I generate the actions of each player according to the corresponding equilibrium strategies. Since the estimation of the number of equilibria is exactly the same as in Chapter 2 through a sequential test of the rank of a matrix, here in Chapter 3 I do not provide evidence of the estimation of the number of the equilibria. Regarding the estimation of the CCPs for different market types and equilibria, we can see that the estimation through matrix decomposition are consistent as sample sizes increase.

3.5 Conclusion

I have studied the identification in finite action games with incomplete information allowing the presence of multiple equilibria and unobserved heterogeneity. In particular, I show that it is possible to identify the cardinality of the combination of multiple equilibria and unobserved market types when there are enough options that players can choose from. However, without strong assumptions, equilibrium strategies can only be identified up to ordering. Consequently, it is impossible to uniquely recover payoffs. I also provide an alternative set identification using the likelihood function.

Understanding what conditions allow us to uniquely recover individual payoffs in static games with the presence of multiple equilibria and unobserved heterogeneity. Even though this paper mainly provides partial identification, it does suggest a direction for future research for pursuing point identification. On the other hand, set identification still provides us information on the payoffs.

Chapter 4

Identification of Dynamic Games with Multiple Equilibria and Unobserved Heterogeneity

4.1 Introduction

Estimation of dynamic discrete games has become one of the fastest growing areas of empirical industrial organization[1]. Three challenges are common in estimation of dynamic games: (1) the computational burden of solving for the fixed points and the curse of dimension in the state space; (2) the treatment of heterogeneity in firm and market characteristics; (3) the present of multiplicity of equilibria. The first challenge dues to the fact estimation requires solving for the equilibrium for each possible value of the parameter. To

[1]see Nevo and Aguirregabiria (2010) for a survey Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games

address this computational problem, researchers proposed two-step estimation approaches which are computational light and easy to implement. The two-step estimators pioneered by Hotz and Miller (1993) significantly broaden the dynamic game application scope.

Addressing unobserved heterogeneity in estimation is important and difficult. Though it is a convenient assumption to assume away the existence of unobserved heterogeneity, it is unrealistic for most applications in empirical IO and also problematic to explain micro data. Not accounting for this heterogeneity can generate significant biases in parameter estimates,[2], thus misleading in understanding of strategic interaction between firms. Accounting for unobserved heterogeneity, and therefore dynamic selection, is also important in dynamic discrete choice models in labor economics. Meanwhile, addressing unobserved heterogeneity is also important in empirical estimation since it is widely incorporated in empirical estimations. It is difficult to cope with unobserved state variables in games because multiple equilibria might exist.

Meanwhile, multiple equilibria are a prevalent feature in dynamic games. Even though the present of multiplicity of equilibria does not necessarily preclude the identification of the dynamic framework (Jovanovic (1989)), it is still unclear under what conditions the identification is obtained when multiple equilibria actually employed in the data. In the dynamic setting, identification and estimation is able to get around the multiplicity concerns

[2]For instance, in the empirical application in Aguirregabiria and Mira (2007), the estimation without unobserved market heterogeneity implies estimates of strategic interaction between firms (i.e., competition effects) that are close to zero or even have the opposite sign to the one expected under competition. While including unobserved heterogeneity in the models results in estimates that show significant and strong competition effects

when Markov Perfect equilibria are considered. This Markovian assumption guarantees that a single time series is generated by only one equilibrium. However, when cross-sectional data is pooled for estimation, one has to assume additionally that the same equilibrium is played in every path observed. The ad hoc equilibrium assumption might result in mis-specification, thus inconsistent estimation of the structural parameters. Moreover, even if the game primitives can be consistently estimated under some assumptions, it is impossible to infer policy effects without the information of the equilibrium selection.

To avoid mis-specification while enabling counterfactual analysis, this paper proposes a methodology to nonparametrically identify dynamic games with incomplete information while considering unobserved market-level heterogeneity (market-type) and multiple equilibria. Assuming that the supports of unobserved market-type and equilibria are discrete and finite, the observable distribution from pooling information from cross-sectional markets results in a mixture structure. Thus, identification follows results developed in Hu and Shum (2012), which utilizes the Markov property of observed and unobserved state variables in dynamic models.

The methodology of identification imposes no restrictions on the cardinality of the equilibrium set or the equilibrium selection rules. Identification proceeds in the following steps. First I identify the cardinality of a new latent variable which combines information of both the unobserved market-type and the multiple equilibria. Second I identify the law of transition for the Markov process. Thus the equilibrium specific conditional choice probability and the transition function for both observed and unobserved state variables are identified. Third, the payoff function are nonparametrically identified with exclusion restrictions as in Pesendorfer and Schmidt-Dengler (2008) for each value of the new latent factor. Consequently, one can distinguish between multiple equilibria and unobserved-market types from comparing the payoff functions. Specifically, multiple equilibria map with the same payoff functions while unobserved-market types are associated with different level of payoffs. As a byproduct, the equilibrium selection and the marginal distribution of the market-type can be identified.

As far as I know, this is the first paper provides identification for dynamic games while incorporates both unobserved heterogeneity and multiple equilibria. The identification result presented in this paper is of real practical importance. This paper provides conditions under which the underlying data generating process can be recovered. One important feature of the identification is that all aspects of the game can be uniquely recovered. Crucially, as long as the conditions provided here are satisfied, consistent estimation is possible regardless the estimation is parametric, semiparametric or nonparametric. This opens up many different avenues for the construction of estimators. At the most parametric level, it assures us that fully parametric models are identified and likelihood inference can proceed in the usual fashion. The result also assumes us of consistent and efficient estimation from semiparametric profile approaches.

This paper contributes to the literature on estimation of dynamic games. For instance, Aguirregabiria and Mira (2007) provides sequential estimator for similar dynamic games and allows for market level unobserved heterogeneity. The market-type is time-invariant, and it does not enter the transition process of the observed state variables, so the transition probability function can still be estimated from transition data without solving the model. Arcidiacono and Miller (2011) consider a more general framework that includes unobserved heterogeneity which can vary over time according to Markov chain process and that can enter both in the payoff function and in the transition of the state variables.

This paper also relates to the literature of identification and estimation of games with multiple equilibria. Various approaches are proposed to tackle multiple equilibria issue. Some researchers assume that the data is generated by the same equilibrium so that the distributions estimated from the data satisfies the equilibrium conditions, e.g. Sweeting (2011). Another strand of literature use bound estimation instead of point estimation, relying on inequalities created by multiple equilibria, e.g., Ciliberto and Tamer (2009). Bajari, Hong, and Ryan (2010) incorporate a parameterized equilibrium selection function into the problem, and Aguirregabiria and Mira (2013) considers static games with both multiple equilibria and payoff relevant heterogeneity together. See De Paula (2012) for a survey of the recent literature on the econometric analysis of games with multiplicity. 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.

This paper is related to current literature on identification with unobserved heterogeneity. In particular, 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 are sufficient for identification in this paper. The identification is close to the identification of dynamic models in Hu and Shum (2012) with continuous unobserved state variables, which are allowed to vary over time. The identification is obtained also 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. Further, the number of equilibria is endogenous in games and needed to be identified first.

The remainder of the paper is organized as follows. I begin with describing the game framework in section 2. Section 3 provides the nonparametric identification of the game. Next section 4 proposes semiparametric estimation following the constructive identification procedure. Monte Carlo evidence for illustrating the finite sample property is provided in section 5. Lastly, section 6 concludes. The Appendix contains the proofs, the figures and the tables.

4.2 Dynamic Game

Consider a model of discrete time, infinite-horizon games with N players.[3] At the beginning of each period $t(t \in \{0, 1, ..., \infty\})$, player i ($i \in \{1, .., N\}$ receives her own private profit shock ϵ_{it} before she decides which action to take, and the action is denoted as a_{it}, where $a_{it} \in A_i = \{0, 1, ..., K\}$. a_t denotes the action vector for all players in period t. Denote the market and individual firm characteristics in period t as x_t. For the equilibrium characterization, whether the state x_t contains the previous action or not does not matter, but it matters for identification. Let ϵ_t represent the private information for all players, i.e $\epsilon_t \equiv (\epsilon_{1t}, ..., \epsilon_{Nt})$. The payoff for player i from choosing action a_{it} while her rivals choosing actions a_{-it} in period t is assumed to be additive separable as follows:

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

[3]See a similar framework used in Ericson and Pakes (1995), Bajari, Chernozhukov, Hong, and Nekipelov (2009), Aguirregabiria and Mira (2007), Pakes, Ostrovsky, and Berry (2007), Kasahara and Shimotsu (2012), Bajari, Benkard, and Levin (2007), Egesdal, Lai, and Su (2013), Beresteanu, Ellickson, and Misra (2010), Ryan (2012) and Pesendorfer and Schmidt-Dengler (2008).

Assume that the state s_t follows a stationary first-order Markov process with a transition function denoted as $g(s_{t+1}|s_t, a_t)$, which is common knowledge among all the players. Consider only Markov Perfect Equilibrium, 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. Since it is a stationary game, let's suppress the t for ease of notation. Again I use $\sigma_i(s, \epsilon_i)$ to denote the probability of choosing the action a_i by player i as $\sigma_i(a_i|s, \epsilon_i)$ given state variable s. Each period, player i's problem is to maximize her own lifetime expected utility discounting by β. Let $W_i(s, \epsilon_i; \sigma)$ be player i's value function given both public state s and her own private information ϵ_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. To define the Markov Perfect Equilibrium using choice probabilities $\{\sigma_i(a_i|s\}_i\}$, first I define the choice specific value function as:

$$V_i(a_i, s) = \Pi_i(a_i, s) + \beta E W_i(s', \epsilon_i'; \sigma)$$

Similar to the choice specific utility in static games, I define the choice specific value function as the determinant part of the lifetime value from choosing that action, excluding the additive profit shocks $\epsilon_i(a_i)$. Given common knowledge s and player's private information ϵ_i, player i is going to choose the action that gives her the highest expected lifetime utility. Notice that for any set of decision rules, the choice specific value, lifetime value and current expected utilities depend on player and her rivals' strategies only through the probability distributions for each option. As in Milgrom and Weber (1985), I also define the Markov

Perfect Equilibrium in terms of choice probabilities in the probability space instead of decision rules as follows.

Definition 4.1. *(**MPE**) A Markov Perfect Equilibrium (MPE) is a collection of CCPs $\{\sigma_i(a_i, s)\}_i$ such that for all player i and all public information 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

$$\begin{aligned} V_i(a_i, s) &= \Pi_i(a_i, s) + \beta E W_i(s', \epsilon_i'; \sigma) \\ W_i(s, \epsilon_i; \sigma) &= \max_{a_i \in A_i} \{V_i(a_i, s) + \epsilon_i(a_i)\} \end{aligned}$$

The utilization of MPE indicates that players' actions are fully determined by the current vector of state variables and own private information. Intuitively, whenever a player observes the same public information, she makes the same decision. History information of the game up to period t does not influence players' decisions.

4.3 Nonparametric Identification Results

This section proposes an identification methodology for the dynamic game while allowing the potential presence of multiple equilibria. Similar to static game, the presence of multiple equilibria is a feature inherent in dynamic games[4]. 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

[4]Pesendorfer and Takahashi (2012) propose several statistical tests to examine multiplicity of equilibria in a similar setup.

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 estimation. Secondly, relying on information from an individual market over time requires a long period of data, limiting application of the estimation. Finally, without information of equilibrium selection, it is difficult to conduct counterfactual analysis. As a result, this section presents identification of dynamic games using cross-section markets information, allowing the existence of multiple equilibria.

Suppose there are N players playing an infinite horizon dynamic game in markets $m = 1, 2, ...M$. Let a_t denote the action vector that the N players choose in period t. The observed market characteristics which I denoted as $x_t \in X$ includes both market and all individual firms' characteristics in period t. Here I assume X has a finite support. As shown in Haller and Lagunoff (2000), stochastic dynamic games also have a finite number of equilibria. I index the equilibrium as $e^* \in 1, 2, ..., Q$, where Q is the total number of equilibria. Note that the number of equilibria does not vary with observables. This is because the equilibrium is a contingent plan for each player in every possible circumstance. In addition to allow for multiple equilibria, I also incorporate a time-variant unobserved market-level heterogeneity, denoted as η_t, $\eta_t \in \Psi \equiv \{\eta^1, ..., \eta^L\}$ also has a finite support. This assumption is common when dealing with permanent unobserved heterogeneity in dynamic structural models. The discrete support of the unobservable implies that the contribution of a market to the likelihood (or pseudo likelihood) function is a finite mixture of likelihoods under the different possible best responses that we would have for each possible market type.

Given the MPE restriction, players make their decisions according to current period's information, i.e. s_t and η_t. History information is irrelevant in the decision making process. 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. Within this setup, players' current decision depends on what they chose in the last period. Similar to most literature, this paper assumes players current decisions a_t depend on their previous actions a_{t-1}, i.e., $s_t = \{x_t, a_{t-1}\}$.

Assumption 4.2. *The market observable x_t and unobservable η_t evolves according to the following:*

(i). $\Pr(\eta_t|x_{t-1}, \eta_{t-1}, a_{t-1}, \Omega_{<t-1}) = \Pr(\eta_t|\eta_{t-1}, x_{t-1}, a_{t-1})$

(ii). $\Pr(x_t|\eta_t, x_{t-1}, \eta_{t-1}, a_{t-1}, \Omega_{<t-1}) = \Pr(x_t|\eta_t, x_{t-1}, a_{t-1})$

where $\Omega_{<t-1} \equiv \{x_{t-2}, \eta_{t-2}, a_{t-2}, ..., x_1, \eta_1, a_1\}$*, the history up to (but not including)* $t-1$

This assumption assumes a Markovian property for players' common knowledge. Moreover, Assumption 1(ii) is a "limited feedback" assumption, which rules out direct feedback from the last period's unobservable η_{t-1}, on the current value of observable x_t. However, it allows indirect effect of η_{t-1} through x_{t-1} and a_{t-1}. Implicitly, this evolution process imposes timing restriction on the game characteristics, which is the unobserved characteristics η_t realized before the observed characteristics x_t. As a result, x_t depends on η_t instead of η_{t-1}. This limited feedback assumption is less restrictive than the assumption made in many applied settings so that the observable x_t evolves independently from the unobservable of any periods so that the state transition of observables can be estimated directly from the data. However, this assumption does rule out the scenario that the alternative timing occurs. The limited feedback assumption is trivial when the unobserved heterogeneity does not vary overtime.

Lemma 4.3. *In a given market, observables and unobservables satisfy the following property, and the joint distribution of observables satisfy the following representations:*

(i). $\{w_t, \eta_t\} \equiv \{a_t, x_t, \eta_t\}$ *follows a stationary first-order Markov process.*

(ii). $\Pr(w_{t+2}, w_{t+1}, w_t) = \sum_{\eta_{t+1}} \Pr(w_{t+2}|w_{t+1}, \eta_{t+1}) \Pr(w_{t+1}, w_t|\eta_{t+1}) \Pr(\eta_{t+1})$

(iii). $\Pr(w_{t+3}, w_{t+2}, w_{t+1}, w_t) = \sum_{\eta_{t+2}} \Pr(w_{t+3}|w_{t+2}, \eta_{t+2}) \Pr(w_{t+2}|w_{t+1}, \eta_{t+2}) \Pr(w_{t+1}, w_t, \eta_{t+2})$

Proof See appendix $\square$

Note that one cannot rule out multiplicity of equilibria in dynamic game frameworks. When data from cross-sectional markets is pooled, it represents a mixture of information from different equilibria. Let e^* as an index of the equilibrium, and $\tau_{t+1} \equiv \{\eta_{t+1}, e^*\}$ captures all the information in both η_{t+1} and the equilibrium. As a result, τ_{t+1} has a finite support. With three periods of data, the joint distribution estimated from pooling data of all markets can be represented as the following:

$$\begin{aligned}
\Pr(w_{t+2}, w_{t+1}, w_t) &= \sum_{e^*} \Pr(w_{t+2}, w_{t+1}, w_t|e^*) \Pr(e^*) \\
&= \sum_{e^*} \sum_{\eta_{t+1}} \Pr(w_{t+2}|w_{t+1}, \eta_{t+1}, e^*) \Pr(w_{t+1}, w_t|\eta_{t+1}, e^*) \Pr(\eta_{t+1}, e^*) \\
&= \sum_{\tau_{t+1}} \Pr(w_{t+2}|w_{t+1}, \tau_{t+1}) \Pr(w_{t+1}, w_t|\tau_{t+1}) \Pr(\tau_{t+1})
\end{aligned}$$

To identify the cardinality of τ_{t+1}, I first introduce following matrix notation while fixing w_{t+1} as $\bar{w}_{t+1}$.

$$
\begin{aligned}
F_{w_{t+2},\bar{w}_{t+1},w_t} &\equiv \left[\Pr\left(w_{t+2}=k,\bar{w}_{t+1},w_t=j\right)\right]_{k,j}, \\
A_{w_{t+2}|\bar{w}_{t+1},\tau_{t+1}} &\equiv \left[\Pr\left(w_{t+2}=k|\bar{w}_{t+1},\tau_{t+1}=q\right)\right]_{k,q} \\
B_{\bar{w}_{t+1},w_t|\tau_{t+1}} &\equiv \left[\Pr\left(\bar{w}_{t+1},w_t=k|\tau_{t+1}=q\right)\right]_{q,k} \\
D_{\tau_{t+1}} &\equiv \operatorname{diag}\Big[\Pr(\tau_{t+1}=1)\ \ \dots\ \ \Pr(\tau_{t+1}=Q\times L)\Big].
\end{aligned}
$$

Those matrices stack the distributions with all possible values that w_t and τ_{t+1} can take. In particular, matrix $F_{w_{t+2},\bar{w}_{t+1},w_t}$ consists of the whole joint distributions of w_{t+2} and w_t, which can be estimated from data. $D_{\tau_{t+1}}$ is a diagonal matrix with the marginal distribution of τ_{t+1} as the diagonal elements, while matrix $A_{w_{t+2}|\bar{w}_{t+1},\tau_{t+1}}$ collects transition probabilities that we need to recover the law of motion.

With above matrix notation, I can rewrite the equation linking observable joint distribution with unknowns in the following matrix representation:

$$
F_{w_{t+2},\bar{w}_{t+1},w_t} = A_{w_{t+2}|\bar{w}_{t+1},\tau_{t+1}} D_{\tau_{t+1}} A_{\bar{w}_{t+1},w_t|\tau_{t+1}}
$$

This equation in a matrix form enables me to identify the number of equilibria using rank inequality, which I stated in the following lemma.

Lemma 4.4. *The rank of the observed matrix $F_{w_{t+2},\overline{w}_{t+1},w_t}$ serves as the lower bound for the number of equilibria, i.e., $Q \geq Rank(F_{w_{t+2},\overline{w}_{t+1},w_t})$, Furthermore, the number of equilibria is identified, in particular, $Q = Rank(F_{w_{t+2},\overline{w}_{t+1},s_t})$ if the following conditions are satisfied (1) $\|X \times A^n\| > Q \times L$ (2) both matrices $A_{w_{t+2}|\overline{w}_{t+1},\tau_{t+1}}$ and $A_{\overline{w}_{t+1},w_t|\tau_{t+1}}$ have full rank (3) all $\Pr(\tau_{t+1})$ are positive*

Proof The proof is the same as lemma 2.3 □

There are several advantages in dynamic games compared to static framework in the identification of the cardinality of latent factors. 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 dynamic games, the full rank condition is required for fixing value of w_{t+1}, which again is easily satisfied. Moreover, with more variation in the measurement, the full rank condition is easily satisfied too.

With four periods of data, the joint distribution of the observables becomes:

$$\begin{aligned} & \Pr(w_{t+3}, w_{t+2}, w_{t+1}, w_t) && (4.1) \\ = & \sum_{e^*} \Pr(w_{t+3}, w_{t+2}, w_{t+1}, w_t | e^*) \Pr(e^*) \\ = & \sum_{e^*} \sum_{\eta_{t+2}} \Pr(w_{t+3} | w_{t+2}, \eta_{t+2}, e^*) \Pr(w_{t+2} | w_{t+1}, \eta_{t+2}, e^*) \Pr(w_{t+1}, w_t, \eta_{t+2}, e^*) \\ = & \sum_{\tau_{t+2}} \Pr(w_{t+3} | w_{t+2}, \tau_{t+2}) \Pr(w_{t+2} | w_{t+1}, \tau_{t+2}) \Pr(w_{t+1}, w_t, \tau_{t+2}) && (4.2) \end{aligned}$$

With the cardinality of the latent factor τ_{t+2} identified, I partition the state space of both w_t and w_{t+3} into a dimension of $Q \times L$ with the criteria of the resulting matrices are full rank. With a little abuse of notation, I denote the new partitioned variables as w_t and w_{t+3} respectively. Fixing w_{t+2} and w_{t+1}, I rewrite equation A.3 into the following matrix expression:

$$F_{w_{t+3}, w_{t+2}, w_{t+1}, w_t} = A_{w_{t+3} | w_{t+2}, \tau_{t+2}} D_{w_{t+2} | w_{t+1}, \tau_{t+2}} B_{w_{t+1}, w_t, \tau_{t+2}} \quad (4.3)$$

Identification of $\Pr(w_{t+3}|w_{t+2}, \tau_{t+2})$ is obtained by 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. After algebra manipulation, I can form a eigenvalue-eigenvector decomposition representation between the observed and unknowns that we are interested. To guarantee unique decomposition, I first state the assumption.

Assumption 4.5. ***(Distinctive Eigenvalues)*** *There exist* $\{w_{t+2}, \bar{w}_{t+2}, w_{t+1}, \bar{w}_{t+1}\}$*, such that*

(1). $\Pr(\bar{w}_{t+2}|w_{t+1}, \tau_{t+2}) \Pr(w_{t+2}|\bar{w}_{t+1}, \tau_{t+2}) > 0$ *for all* τ_{t+2}

(2). $C(w_{t+2}, \bar{w}_{t+2}, w_{t+1}, \bar{w}_{t+1}|\tau_{t+2} = i) \neq C(w_{t+2}, \bar{w}_{t+2}, w_{t+1}, \bar{w}_{t+1}|\tau_{t+2} = j)$ *for any* τ_{t+2} $i \neq j$*, where*

$$C(w_{t+2}, \bar{w}_{t+2}, w_{t+1}, \bar{w}_{t+1}|\tau_{t+2}) = \frac{Pr(w_{t+2}|w_{t+1}, \tau_{t+2})Pr(\bar{w}_{t+2}|\bar{w}_{t+1}, \tau_{t+2})}{Pr(\bar{w}_{t+2}|w_{t+1}, \tau_{t+2})Pr(w_{t+2}|\bar{w}_{t+1}, \tau_{t+2})}$$

The distinctive eigenvalues assumption is empirically testable because the matrix for the eigen-decomposition can be computed from the data. Moreover, the eigenvectors $A_{w_{t+3}|w_{t+2},\tau_{t+2}}$ 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.

Lemma 4.6. ***(Identification of*** $\Pr(w_{t+3}|w_{t+2}, \tau_{t+2})$*): With four periods of data, assumption 1 and conditions in lemma 1 satisfied, for each possible value of* w_{t+2}*,* $\Pr(w_{t+3}|w_{t+2}, \tau_{t+2})$ *is uniquely identified up to ordering of* τ_{t+2}*.*

Proof See Appendix □

Note that with the eigenvalue-eigenvector decomposition representation for each w_{t+2}, the eigenvalues are identified up to ordering of τ_{t+2}. Since η_{t+2} combines information from both market type and multiplicity of equilibria, the conventional monotonicity property is not appropriate to use to order the eigenvalues. However, for each w_{t+2}, using marginal distribution of $\Pr(w_{t+3}|w_{t+2})$, we can recover the marginal distribution of η_{t+2}. If the probability of η_{t+2} varies with different elements of η_{t+2} can take, then we can match the marginal distribution to preserve the ordering of the eigenvalues when the decomposition is conducted for different values of w_{t+2}. In the following I state this assumption.

Assumption 4.7. ***(Distinctive Marginal Distribution)*** *The marginal distribution of η_{t+2} varies for different value that η_{t+2} can be:* $\Pr(\eta_{t+2} = i) \neq \Pr(\eta_{t+2} = j)$, *where* $i \neq j$

This assumption is also empirically testable. Now I move to show that the law of motion can be identified using four periods of data when $\Pr(w_{t+3}|w_{t+2}, \tau_{t+2})$ can be uniquely recovered.

Lemma 4.8. ***(Markov Law of Motion)*** *With four periods of data, and* $\Pr(w_{t+3}|w_{t+2}, \tau_{t+2})$ *is known, the Markov law of motion* $\Pr(w_{t+3}, \tau_{t+3}|w_{t+2}, \tau_{t+2})$ *can be uniquely identified.*

Proof See Appendix □

Initial condition $\Pr(w_t, \tau_t)$ plays an important role in simulating the game to do estimation while this information is impossible to obtain from the data. However, following lemma states that it can be uniquely recovered as a byproduct of the main identification.

Lemma 4.9. ***(Initial Condition)****: Under assumptions 1,2,3,4...., the initial density distribution* $\Pr(w_t, \tau_t)$ *can be uniquely recovered from four periods of data.*

Proof See Appendix □

As a byproduct of the initial joint distribution of observables w_t and unobservables τ_t, the marginal distribution of uobservables can be identified. The equilibrium selection therefore can be identified once we can distinguish between unobserved types and multiple equilibria.

With the law of motion, the conditional choice probability and transition of observed and unobserved factors can all be identified by taking marginal with respect to the law of motion. Below I first provide that the payoff function can be identified with CCPs and transition functions known and stated the result in the following lemma. Then I will discuss how I can distinguish whether two τ are two equilibria associated with the same market-type or two different market-type. To nonparametrically identify payoff functions, the following two assumptions are necessary, and they are standard assumptions imposed in the existing literature.

Assumption 4.10. ***(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.

Assumption 4.11. ***(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. $\pi_i(a_i = k, a_{-i}, s) \equiv \pi_i(a_i = k, a_{-i}, s_i)$.*

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

Lemma 4.12. ***(Identification of payoff functions)*** *Under assumptions 4.10 and 4.11 and all the conditional choice probability and state transition is known, payoff functions are identified nonparametrically.*

Proof See appendix □

With everything is known, the identification of payoff functions is back to the traditional case with a unique equilibrium (see Bajari, Chernozhukov, Hong, and Nekipelov (2009)). 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.

Theorem 4.13. ***(Identification of Dynamic Games with Incomplete Information)*** *With the conditions in lemma 4 satisfied, assumptions 1-5, the cardinality and initial marginal distribution of the unobserved heterogeneity , the number of equilibria Q, the equilibrium selection, the strategies of each player in each equilibrium and the payoff function are nonparametrically identified in dynamic games with four periods of data.*

Given that all the conditional choice probabilities, state transition and initial conditions can be uniquely identified, I can identify payoffs for different values of τ with exclusion restrictions. To distinguish market-level unobserved heterogeneity from multiple equilibria, one just needs to compare payoffs from any different values of τ. If the payoffs are the same for different values of τ, then the two τs are two different equilibria associated with the same market-type. If the payoffs are different for the two different values of τ, then the payoffs are associated with different market-types. Equivalently, we can divide all the τ into different groups in which contains different equilibria belongs to the same market-type. Additionally, the equilibrium selection mechanism can be identified through the marginal distribution of τ, which just a conditional distribution within each group. Moreover, the cardinality of the unobserved heterogeneity is also identified as the number of distinct groups.

4.4 Estimation

This section provides semi-parametric estimation for the dynamic game. The estimation follows exactly the identification procedure. Specifically, first the cardinality of the newly created latent variable is estimated nonparametrically by estimation of the rank of the matrix constructed by observable joint distribution. Then each equilibrium strategies and equilibrium specific transition are estimated through eigenvalue-eigenvector decomposition. Last, the structural parameters are estimated through minimizing distance between the equilibrium strategies and its best response.

Estimation of the cardinality The cardinality of the latent factor τ can be estimated via estimating the rank of the matrix constructed by observable joint distribution, i.e. $Q = rank(F_{w_{t+2},\bar{w}_{t+1},w_t})$, where the ij_{th} element of the matrix can be estimated using simple frequency such as: $\Pr(w_{t+2} = w^i, \bar{w}_{t+1}, w_t = w^j) = \frac{I(w_{t+2}=w^i,\bar{w}_{t+1},w_t=w^j)}{N}$. Provided the stationary environment, one can tailor the panel data into a data in which each observation consists three periods of information, which increases the estimation preciseness.

A sequence of tests is performed to estimate the rank of the matrix constructed by joint distribution among three periods of observables. There is a lot of ongoing research on testing rank of a matrix through the estimates of that matrix, e.g., characteristic roots of a quadratic form built from the matrix in Robin and Smith (2000). Those testing methodologies are proposed for general matrices. Since the matrices used here need to be square and invertible, and the determinant of an invertible matrix is non zero. Thus I am going to use the determinant of the partitioned of the corresponding matrix as the statistic for the testing.

The testing procedure is as follows. The matrix $\hat{F}_{w_3,\bar{w}_2,w_1}$ estimated through simple frequency has a dimension of $l * l$. Let $\hat{F}^J_{w_3,\bar{w}_2,w_1}$ denote the matrix with dimension $J \times J$ through randomly partitioning the original space of w_1 and w_3 into a new space with J number of support. The sequential procedure is given as follows: the sequence of hypotheses $H^J_0 : det(F^J) = 0$ is tested against the alternatives $H^J_1 : det(F^J) \neq 0$ in decreasing order. If no rejection occurs until $J = 1$, then there is a unique equilibrium. 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 α 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 α 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 α_N goes to zero as the sample size N goes to infinity but not faster than a given rate. As in Hosoya (1989), if α_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 estimator provided by the sequential testing procedure will converge in probability to the true rank.

Estimation of the CCPs With the cardinality of the latent variable is known, the estimation of all CCPs exactly follows the identification procedure. First of all, $\Pr(w_{t+3}|w_{t+2}, \tau_{t+2})$ can be estimated as an eigenvectors of the matrix decomposition following the identification procedure as in lemma 3. Then the markov law of motion can be estimated from matrix manipulation following lemma 4. Consequently, the conditional choice probability $\Pr(a_{it}|x_t, a_{t-1}, \eta_t)$ and the transition function can be estimated nonparametrically.

Parametric Estimation of the 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 $\pi_i(a_i, a_{-i}, x) = \pi_i(a_i, a_{-i}, x; \theta)$, and suppose market characteristics are discrete with dimension of d, $x \in \{X_1, ..., X_d\}$.

Pioneered by Hotz and Miller (1993)[5], 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 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

[5]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.

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 $\tilde{\sigma}(a_t|x_t, \eta_t, a_{t-1})$. The equilibrium condition is represented by a general mapping denoted as $h(\sigma, \theta) = \sigma(a_t|x_t, \eta_t, a_{t-1}) - \Gamma(a_t|x_t, \eta_t, a_{t-1}); \theta) = 0$, which holds for every x. The least squares estimator estimates the parameters of interest by forcing the constraints:

$$h(\hat{\sigma}, \theta) = \hat{\sigma}(a_t|x_t, \eta_t, a_{t-1}) - \Gamma(\hat{\sigma}(a_t|x_t, \eta_t, a_{t-1}); \theta) = 0$$

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 $\tilde{\sigma}_M$ 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 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

$$\hat{\theta}(W_M) = argmin_\theta \quad [\hat{\sigma} - \Gamma(\hat{\sigma}; \theta)]' \, W_M \, [\hat{\sigma} - \Gamma(\hat{\sigma}; \theta)]$$

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.

The structural parameters are consistent and asymptotically normal with regularity conditions and consistency and asymptotical normality of the CCPs and transition function estimated in the first-step. If the cardinality of the latent variable is known, CCPs and transition function are estimated through matrix algebra following eigenvalue-eigenvector decomposition. Given that eigenvalues and eigenvectors can be represented as an analytic expression from the elements of the associated matrix (Andrew, Chu, and Lancaster (1993)), CCPs and transition function are consistently estimated and are asymptotical normal in the first step. Therefore, the least square estimators for the structural parameters are consistently estimated and asymptotical normal with the regularity conditions provided in Pesendorfer and Schmidt-Dengler (2008)

Hypothesis testing The fact that we can identify the parameter for each possible value of the latent variable $\tau = \{\eta, e^*\}$ provides a testable implications for us to distinguish whether two values of τ are just two equilibria associated with the same market type, or two different market types. For notation simplicity purpose, I denote the two θ as θ^1 and θ^2 and the estimators as $\hat{\theta}^1$ and $\hat{\theta}^2$. Specifically, if we conduct the following hypothesis:

$$H_0 : \theta_0^1 = \theta_0^2$$

against the alternative

$$H_1 : \theta_0^1 \neq \theta_0^2$$

then the two τs belong to the same market-type if we fail to reject the null hypothesis

H_0, otherwise we conclude that the two τs represent two different market types. The asymptotical properties of the two estimators are as follows:

$$\frac{1}{M}(\hat{\theta}^i_M - \theta^i_0) \rightarrow_d N(0, \Sigma_i) \qquad i = 1, 2$$

Thus, under the null, I construct the following statistic:

$$T = M(\hat{\theta}^1_M - \hat{\theta}^2_M)^T \widehat{var(\hat{\theta}^1_M - \hat{\theta}^2_M)}^{-1} (\hat{\theta}^1_M - \hat{\theta}^2_M) \rightarrow_d \chi^2(k)$$

where $\widehat{var(\hat{\theta}^1_M - \hat{\theta}^2_M)}$ is a estimate of the covariance of $\hat{\theta}^1_M - \hat{\theta}^2_M$, which in practical can be computed through bootstrap.

4.5 Monte Carlo Evidence

This section investigates the finite sample property of the proposed simple estimators in a Monte Carlo study. To make thing simple, I consider a simple dynamic oligopoly game with multiple equilibria. The methodology proposed in this paper nests this simple example. The game was illustrated and analyzed in more detail in Pesendorfer and Schmidt-Dengler (2008). The same game is also analyzed in Pesendorfer and Takahashi (2012).

Consider a setting with two players, binary actions $\{0, 1\}$ and binary states $\{0, 1\}$. Consider an infinite horizon game with two players and choose entry/exit. The specification is as following: the distribution of the profitability shocks F is standard normal. The discount rate is fixed at 0.9. The state transition law is given by $s^{t+1}_i = a^t_i$. Period pay-offs are symmetric and are parameterized as follows:

$$\pi_{a_i,a_j,s_i} = \begin{cases} 0, & \text{if } a_i = 0; s_i = 0; \\ x, & \text{if } a_i = 0; s_i = 1; \\ \pi^1 + c, & \text{if } a_i = 1; a_j = 0; s_i = 0; \\ \pi^2 + c, & \text{if } a_i = 1; a_j = 1; s_i = 0; \\ \pi^1, & \text{if } a_i = 1; a_j = 0; s_i = 1; \\ \pi^2, & \text{if } a_i = 1; a_j = 1; s_i = 1. \end{cases}$$

where $x = 0.1, c = -0.2, \pi^1 = 1.2$ and $\pi^2 = -1.2$. The period payoff can be viewed as switching costs as entry/exit in a dynamic game. A player that selects action 1 receives monopoly profits if she is the only active player, and she receives duopoly profits otherwise. In addition, a player that switches states from 0 to 1 incurs the reactivated c; while a player that switches from 1 to 0 receives the exit value x.

Multiplicity. The game illustrates that multiple equilibria are present. There are five equilibria theoretically but this Monte Carlo illustration focuses on two asymmetric equilibria of the three equilibria described in Pesendorfer and Schmidt-Dengler (2008). In equilibrium (i), player two is more likely to choose action 0 than player one in all states. The ex ante probability vectors for both players are given by: $\sigma(a_1 = 0|s_1, s_2) = (0.27; 0.39; 0.20; 0.25)', \sigma(a_2 = 0|s_1, s_2) = (0.72; 0.78; 0.58; 0.71)'$ where the order of the elements in the probability vectors corresponds to the state vector $(s_1, s_2) = ((00); (01); (10); (11))$. In equilibrium (ii), player two is more likely to choose action 0 than player one in all states with the exception of state (1 0). The probability vectors are given by $\sigma(a_1 = 1|s_1, s_2) = (0.38; 0.69; 0.17; 0.39)', \sigma(a_2 = 0|s_1, s_2) = (0.47; 0.70; 0.16; 0.42)'$.

The Monte Carlo study considers estimation of the game primitives. The simulated data are generated by randomly drawing a time series of actions from the calculated equilibrium choice probabilities described above for each of the equilibria (i)-(ii) respectively. The initial state is taken as the four different states with a equal probability of 0.25. Given that the asymptotic property of the estimators relies on the number of markets goes to infinity and identification requires four periods of data, I varies the number of markets while keep the length of the time series as four during all the simulation. The parameter λ denotes the fraction of markets that adopt equilibrium (i) while $1 - \lambda$ denotes the fraction of markets that adopt equilibrium (ii).

The cardinality is estimated through a sequential tests on the rank of the matrix constructed by joint distributions of four periods of data. From figure C.1, the estimator is consistent since the frequency of accept the null that the cardinality equals to two approaches 1 when the sample size increases. Estimation of law of motion is converging to the true but slowly.

4.6 Conclusion

I have developed a methodology to nonparametrically identify finite action games with incomplete information allowing for the presence of multiple equilibria and unobserved heterogeneity. The main contribution of this paper is to present conditions under which all aspects of the game can be uniquely recovered. Specifically, the cardinality of the overall latent factors can be identified nonparametrically. The law of motion and CCPs which are latent factor variant can also be uniquely recovered. With CCPs and transition functions identified, the payoffs can be nonparametrically identified with exclusion restrictions. Disentangling equilibria and unobserved heterogeneity can be obtained from testing payoffs.

The estimation follows identification step-by-step. For parameterized payoff functions, the structural parameters are proposed to be estimated from minimizing the distance of CCPs and its best response. However, there are other alternatives developed to increase estimation efficiency of the one proposed here in this paper such as sequential estimation by Aguirregabiria and Mira (2007) and Egesdal, Lai, and Su (2013). With the identification conditions satisfied, a future research direction will be to incorporate existing estimation method into the settings to allow for both multiple equilibria and unobserved heterogeneity.

Chapter 5

Conclusion

This dissertation mainly is composed by three papers, studying point identification of static games with multiple equilibria only, set identification of static games with both multiple equilibria and unobserved heterogeneity and point identification of dynamic games with both multiple equilibria and market level unobserved heterogeneity. The identification uses results from measurement error literature by treating either equilibria and/or market latent types as a latent variable.

Chapter 2 of this dissertation tackles the problem of multiple equilibria in static game settings in both cross-sectional and panel data structure. If cross-section data is available, the traditional assumption that private payoff shocks are independent across actions and players plays an important role in identification. When panel data is available, the key condition is that the equilibrium employment evolves according to a first-order Markov process. Identification of payoff is exactly the same as the case with cross-sectional data. Empirical application provides evidence that multiple equilibria do exist in the game where radio stations strategically determine when to air their commercials.

Chapter 3, as an natural extension of Chapter 2, incorporates unobserved market-level heterogeneity into the identification. However, without being able to disentangle equilibria and the payoff relevant-unobserved market types, it is impossible to uniquely recover the payoffs. As a result, Chapter 3 provides partial identification of the parameterized payoff functions.

Chapter 4 is to study the identification in dynamic games where multiple equilibria are prevalent and unobserved heterogeneity is empirically important. With imposing assumption such as the same equilibrium is employed in the data, I show that the cardinality of overall latent factors can be uniquely recovered, so does the law of motion. With CCPs associated with the overall latent factors, payoffs can be nonparametrically identified with exclusion restrictions for each value of the overall latent factor. Consequently, one can distinguish between multiple equilibria and unobserved-market types from comparing the payoff functions. Specifically, multiple equilibria map with the same payoff functions while unobserved-market types are associated with different level of payoffs. As a byproduct, the equilibrium selection and the marginal distribution of the market-type can be identified.

This dissertation mainly focus on identification, and estimation simply follows the constructive identification procedure. However, the nonparametric estimation of equilibrium related components requires a very rich data, which sometimes is not available. Moreover, the payoff estimators are from the minimal distance of CCPs and its best response, which are also not necessarily efficient. Thus, a potential future research avenue will be developing or incorporate existing estimation technique to do estimation and provide inferences.

Multiple equilibria are generic feature in games. Addressing multiplicity of equilibria is important to obtain consistent estimation of the underlying game structures given the widely usage of games to analyze strategic interactions between players. For instance, games where the actions of players are continuous. When the actions of players are continuous, it provides more identification power given that the first step identification comes from the variation of actions. Moreover, when actions are continuous, the full rank conditions might be not that restrictive since now we can reposition the action space with lots of ways. The infinite ways of partition increase the possibility with which we might get one particular partition that full rank condition satisfied.

Appendix A

Appendix to Chapter 2

A Proofs

Proof of Lemma 2.3 Based on conditional independent assumptions, the joint distribution of actions from two players can be expressed as:

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

Rewrite it into a matrix form:

$$F_{a_1,a_2} = A_{a_1|e^*} D A^T_{a_2|e^*}$$

With assumptions that $(K+1)^l > Q$ and full rank of both matrices $A_{a_1|e^*}$ and $A_{a_2|e^*}$, then according to the following inequality regarding the rank of matrix F_{a_1,a_2}:

$$Rank(A_{a_1|e^*}) + Rank(A_{a_2|e^*}) - Q \leq Rank(F_{a_1,a_2}) \leq \min\{Rank(A_{a_1|e^*}), Rank(A_{a_2|e^*})\} \tag{A.1}$$

I conclude that $Rank(F_{a_1,a_2}) = Q$. □

Proof of Lemma 2.4 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 $rank(A) = Q$, 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, $\eta_i(1+\lambda_i) = 0$ implies that $\eta_i = 0$. Then $\eta_k = 0$ for $k = 1, ...Q$. Thus, $a_1, a_2, ..., a_i + a_{Q+1}, a_{i+1}, ..., a_Q$ are linear independent. $\square$

Proof of Lemma 2.6 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

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

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 $D_{e^*} = A^{-1}_{\tilde{a}_1|e^*} F_{\tilde{a}_1}$.

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:

$$F_{a_1,\tilde{a}_2} = A_{a_1|e^*} D_{e^*} A^T_{\tilde{a}_2|e^*}$$

Since D_{e^*} and $A^T_{\tilde{a}_2|e^*}$ are identified and invertible, then $A_{a_1|e^*}$ is identified. $A_{a_2|e^*}$ can be identified in the same procedure. $\square$

Proof of Theorem 2.11 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:

$$\begin{aligned}
\Pr(a_{t+2}, a_{t+1}, a_t) &= \sum_{e^*_t, e^*_{t+1}, e^*_{t+2}} \Pr(a_{t+2}, e^*_{t+2}, a_{t+1}, e^*_{t+1}, a_t, e^*_t) \\
&= \sum_{e^*_t, e^*_{t+1}, e^*_{t+2}} \Pr(a_{t+2}|e^*_{t+2}) \Pr(e^*_{t+2}|e^*_{t+1}) \Pr(a_{t+1}|e^*_{t+1}) \Pr(e^*_{t+1}, e^*_t) \Pr(a_t|e^*_t) \\
&= \sum_{e^*_t, e^*_{t+1}} \Pr(a_{t+2}|e^*_{t+1}) \Pr(a_{t+1}|e^*_{t+1}) \Pr(a_t|e^*_t) \Pr(e^*_{t+1}, e^*_t) \\
&= \sum_{e^*_{t+1}} \Pr(a_{t+2}|e^*_{t+1}) \Pr(a_{t+1}|e^*_{t+1}) \Pr(a_t|e^*_{t+1}) \Pr(e^*_{t+1})
\end{aligned}$$

where $\Pr(a_l|e^*_{t+1})$ represents the probability of the players choosing action a_l 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 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 $\xi_\tau = \{b_{\tau 1}, ..., b_{\tau Q}\}$ for $\tau = t, t+2$ so that the matrix F_{ξ_{t+2},ξ_t} defined accordingly is invertible. Matrix representation for the joint probability distribution equations with the following matrices definitions, we have

$$F_{\xi_{t+2},\xi_t} = A_{\xi_{t+2}|e^*_{t+1}} D A^T_{\xi_t|e^*_{t+1}} \tag{A.2}$$

$$F_{\xi_{t+2},\xi_t,a_{t+1}=k} = A_{\xi_{t+2}|e^*_{t+1}} D_{a_{t+1}=k|e^*_{t+1}} D A^T_{\xi_t|e^*_{t+1}} \tag{A.3}$$

Since matrix F_{ξ_{t+2},ξ_t} is invertible, I can post-multiply $A^{-1}_{\xi_{t+2},\xi_t}$ into both sides of equation A.3, leading to the following main equation.

$$F_{\xi_{t+2},\xi_t,a_{t+1}=k} F^{-1}_{\xi_{t+2},\xi_t} = A_{\xi_{t+2}|e^*_{t+1}} D_{a_{t+1}=k|e^*_{t+1}} A^{-1}_{\xi_{t+2}|e^*_{t+1}} \tag{A.4}$$

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

Assumption A.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 $\Pr(a_{it+1|e^*_{t+1}})$ can be identified by summing over a_{-it+1} on $\Pr(a_{t+1}|e^*_{t+1})$.

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

$$\Pr(\xi_{t+2}|e^*_{t+1}) = \sum_{e^*_{t+2}} \Pr(\xi_{t+2}, e^*_{t+2}|e^*_{t+1}) = \sum_{e^*_{t+2}} \Pr(a_{t+2}|e^*_{t+2}) \Pr(e^*_{t+2}|e^*_{t+1})$$

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 $\Pr(\xi_{t+2}|e^*_{t+2})$ is identified as eigenvectors of the decomposition with normalization by column sum equals to 1, and $\Pr(a_{t+2}|e^*_{t+2})$ is the same as $\Pr(a_{t+1}|e^*_{t+1})$. Rewrite them into matrix form so that the equilibrium evolution process is identified. $\square$

B Asymptotic Properties of the Estimators

This section discusses the consistency and asymptotic normality of estimators. Since relevant proofs are standard in the literature, I just point to the relevant literature for reference. First I present that the rank of a generic matrix estimated from a sequential testing is consistent. Then I provide conditions that equilibrium CCPs and the structural parameters are consistently estimated. And briefly discuss the post-selection inference.

If the number of equilibria is known, then the CCP can be estimated through eigenvalue eigenvector decomposition of the matrix constructed by joint distribution of players. The rest of CCPs can be estimated through matrix algebra. As in Andrew, Chu, and Lancaster (1993), eigenvalues and eigenvectors can be represented as an analytic function of the elements consisting of the matrix for the decomposition. Thus, the CCPs are consistently estimated and asymptotically normal, which I state in the following: there exists a sequence of estimator $\hat{p}_m$ of p such that

$$\hat{p}_M \to p(\theta_0), a.s$$

$$\sqrt{M}(\hat{p}_M - p(\theta_0)) \to_d N(0, \sigma(\theta_0))$$

To prove the asymptotical normality of the structural parameter estimators in the second step, CCPs estimators in the first step needs to be consistent and asymptotical normal distributed(Pesendorfer and Schmidt-Dengler (2008)). With the CCPs of consistency and asymptotical normality, assumptions needed for asymptotical properties of structural estimators are stated in the following:

A1: Θ is a compact set.

A2: the true value θ is in the interior of Θ

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

A4: θ 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 π are twice continuously differentiable in θ.

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.

When the number of equilibria is unkonwn, it can be estimated through a sequential testing. The null hypothesis of the test is that the rank of the corresponding matrix equals to be predetermined number. Note that if the critical value associated with the significant level α 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 α 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 α_N goes to zero as the sample size N goes to infinity but not faster than a given rate. Hosoya (1989) shows that if α_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.

One thing worth noting is that that the estimation of the number of equilibria serves as a model selection procedure. When we obtain the number of equilibria differing from the true one, it is hard to explain the CCPs estimated in the next step. With the model selection, the structural parameters have properties that are different from what is conventionally assumed. To better understand the asymptotical property of the parameters, 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 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))

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 the true structural parameters as Q^0 and θ^0 respectively. Let $\hat{Q}_m$ 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:

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

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

C Graphs and Tables

This section includes all the graphs and tables mentioned in the paper.

Figure A.1: The Best Response Function for Different Values of s

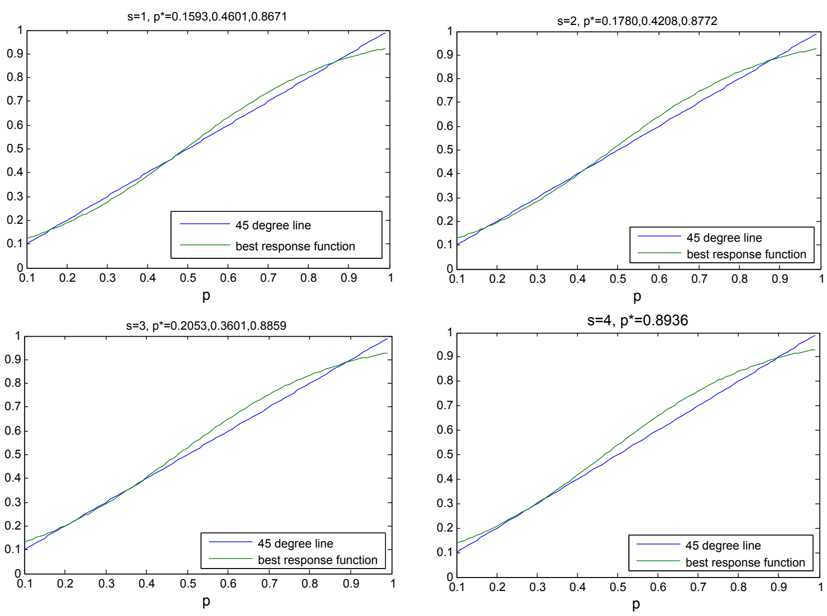

Figure A.2: Frequency of Selecting the True Number of Equilibria

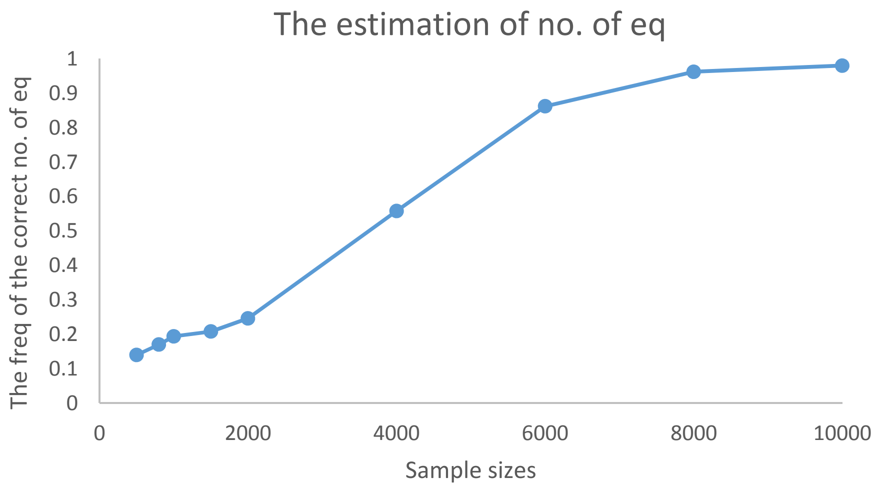

Table A.1: Monte Carlo Evidence: Cross-sectional Data

	x=1		x=2		x=3		x=4	
	DGP	Est	DGP	Est	DGP	Est	DGP	Est
Q_s	2	-	2	-	2	-	1	-
$\sigma(a=0\|s,e^*=1)$	0.1593	0.1594	0.1780	0.1784	0.2053	0.2053	n/a	
		(0.012)		(0.014)		(0.012)		
$\sigma(a=0\|s,e^*=2)$	0.8671	0.8673	0.8772	0.8773	0.8859	0.8857	0.8936	0.8934
		(0.013)		(0.0112)		(0.0213)		(0.004)
$\Pr(e^*)$	0.5	0.5004	0.5	0.4994	0.5	0.4995	0	0
		(0.017)		(0.018)		(0.018)		(0)

[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 A.2: Monte Carlo Evidence: Model Primitives

	DGP	Cross-section		Panel data	
		Unique Eq	Multiple Eq	Unique Eq	Multiple Eq
Strategic Interaction δ	2.5	2.8245	2.5054	2.2762	2.5088
		(0.0405)	(0.0190)	(0.0433)	(0.0444)
Market Effect β	0.04	-0.0236	0.0398	0.0093	0.0395
		(0.0056)	(0.0048)	(0.0071)	(0.0094)

[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 A.3: Monte Carlo Evidence: Panel Data

	x=1		x=2		x=3		x=4	
	DGP	Est	DGP	Est	DGP	Est	DGP	Est
Q_s	2	-	2	-	2	-	1	-
$\sigma(a=0\|s,e^*=1)$	0.1593	0.1605	0.1780	0.1799	0.2053	0.2083	n/a	
		(0.017)		(0.018)		(0.014)		
$\sigma(a=0\|s,e^*=2)$	0.8671	0.8581	0.8772	0.8733	0.8859	0.8765	0.8936	0.8936
		(0.0322)		(0.017)		(0.028)		(0.004)
$\Pr(e^*=2\|e^*=1)$	0.1	0.094	0.25	0.231	0.3	0.287		
		(0.027)		(0.045)		(0.042)		
$\Pr(e^*=2\|e^*=2)$	0.8	0.826	0.85	0.865	0.9	0.931		
		(0.050)		(0.031)		(0.046)		

[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 A.3: Timing Patterns for Commercials across Markets

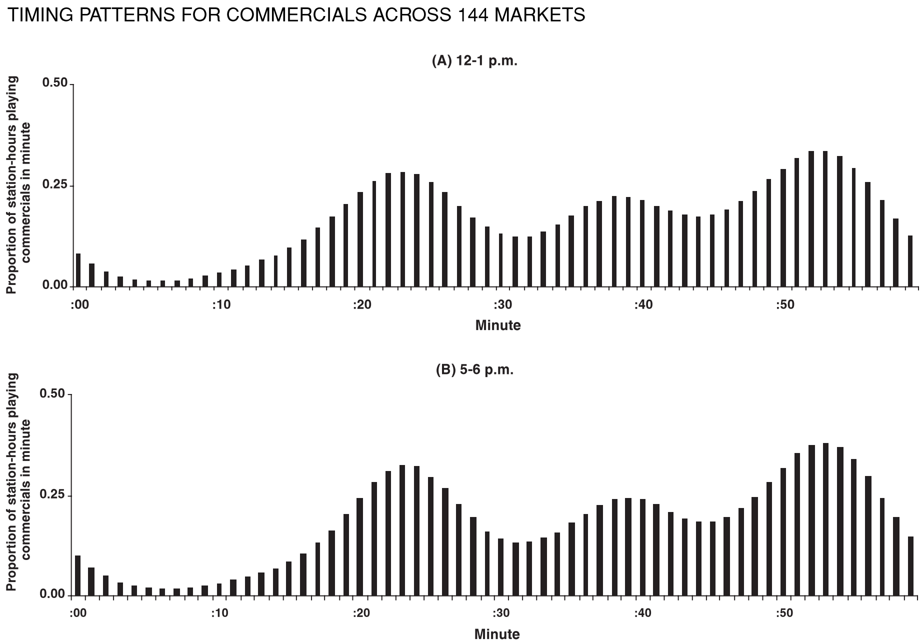

Table A.4: Summary Statistics

Variable	Obs	Mean	std. Dev	Min	Max
No. Players	108554	12.93453	3.174468	2	20
Timing	108554	.4985537	.5000002	0	1
Day	108554	31.42016	17.56653	1	59
Hour	108554	16.46269	3.153961	12	21
Market(big=1)	108554	.5168672	.4997177	0	1

Figure A.4: Timing Patterns for Commercials in Different Markets

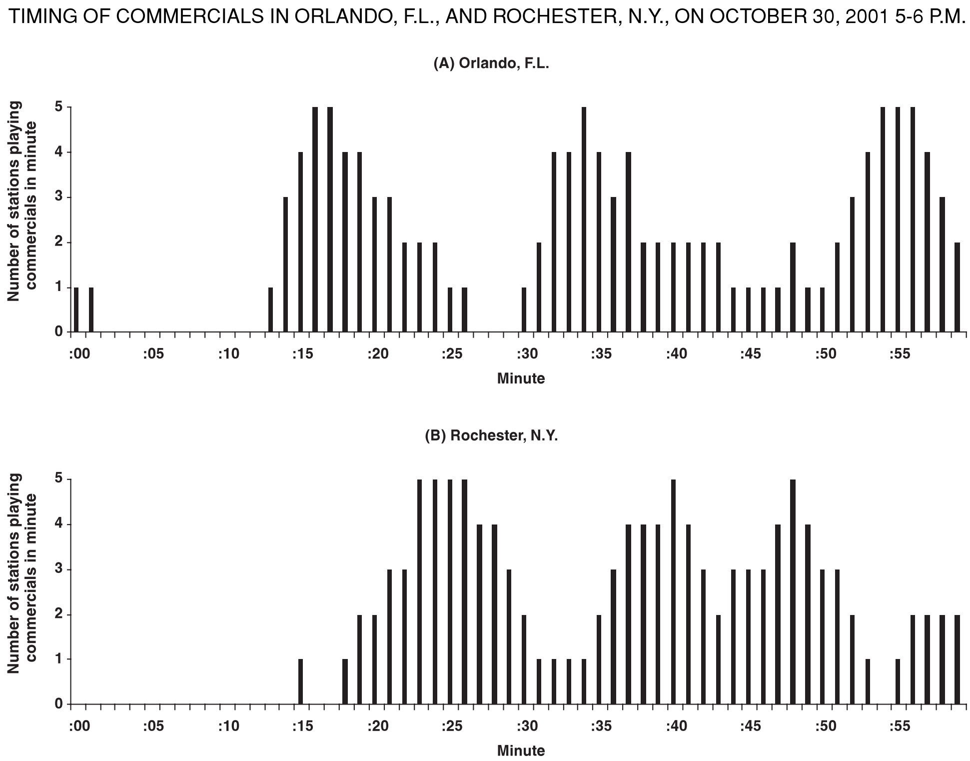

Table A.5: Estimates of Commercial Airing Strategies with Cross-sectional Data

		Market Size		Time			
	All market	Big	Small	Drivetime		Non-drivetime	
				4-5 PM	5-6 PM	12-1 PM	9-10PM
Q_s	2	1	2	2	2	1	1
$\Pr(e^* = 1)$	0.2846	1	0.2732	0.3307	0.4192	1	1
	(0.1293)	-	(0.1230)	(0.1191)	(0.1275)	-	-
$\Pr(a = 0 \vert e^* = 1)$	0.6565	0.5061	0.6582	0.6841	0.6687	0.5203	0.4921
	(0.1875)	(0.0099)	(0.1171)	(0.2206)	(0.2139)	(0.0043)	(0.0043)
$\Pr(a = 0 \vert e^* = 2)$	0.4288	-	0.4287	0.3974	0.3700	-	-
	(0.0630)	-	(0.0498)	(0.0884)	(0.1107)	-	-
α	-0.0055	-0.0242	-0.0056	-0.0092	-0.0051	-0.0811	0.0316
	(0.3258)	(0.0389)	(0.1007)	(0.3019)	(0.2197)	(0.0172)	(0.0172)
δ	2.0520	0	2.0532	2.0736	2.0665	0	0
	(0.3147)	-	(0.1089)	(0.2844)	(0.1999)	-	-

[1] The number in brackets is the standard deviation computed through bootstrap with 500 repetition
[2] For markets with unique equilibrium, δ is assumed to be zero

Table A.6: Estimates of Commercial Airing Strategies with Panel Data

		Market Size		Time			
	All Market	Big	Small	Drivetime		Non-drivetime	
				4-5 PM	5-6 PM	12-1 PM	9-10PM
Q_s	2	1	2	2	2	1	1
$\Pr(e^*_{t+1} = 1 \mid e^*_t = 1)$	1.0000	1	1.0000	1.0000	1.0000	1	1
	(0.0192)	-	(0.4311)	(0.0814)	(0.0485)	-	-
$\Pr(e^*_{t+1} = 2 \mid e^*_t = 2)$	0.8714	-	0.9777	0.8749	0.9239	-	-
	(0.0192)	-	(0.4076)	(0.0959)	(0.0908)	-	-
$\Pr(a = 0 \mid e^* = 1)$	0.5612	0.5116	0.6576	0.5769	0.6434	0.5096	0.4975
	(0.0234)	(0.0085)	(0.1523)	(0.0384)	(0.0414)	(0.0084)	(0.0084)
$\Pr(a = 0 \mid e^* = 2)$	0.3617	-	0.4382	0.3037	0.3785	-	-
	(0.0371)	-	(0.0608)	(0.0719)	(0.0304)	-	-
α	0.0037	-0.0465	-0.0053	0.0108	-0.0022	-0.0383	0.0102
	(0.0152)	(0.0340)	(0.0622)	(0.0195)	(0.0301)	(0.0336)	(0.0338)
δ	2.0402	0	2.0055	2.0860	2.0500	0	0
	(0.0336)	-	(0.0719)	(0.2401)	(0.0335)	-	-

[1] The number in brackets is the standard deviation computed through bootstrap with replacement, with 500 repetition
[2] For markets with unique equilibrium, β is assumed to be zero

Appendix B

Appendix to Chapter 3

A Graphs and Tables

Figure B.1: The CCPs associated with different unobserved heterogeneity

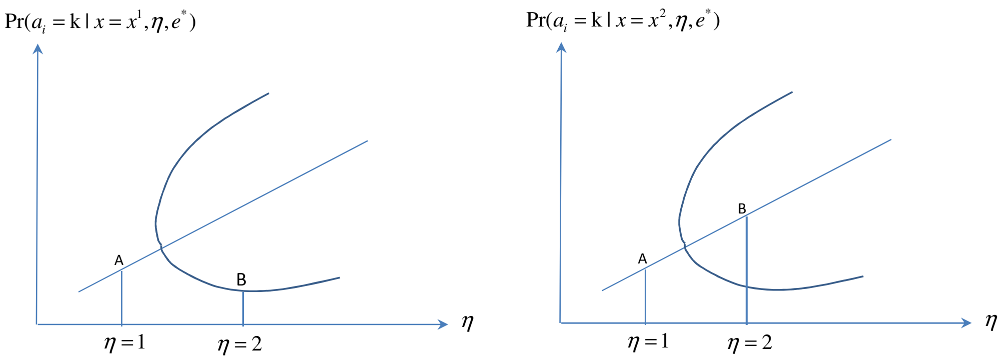

Figure B.2: The CCPs associated with different unobserved heterogeneity

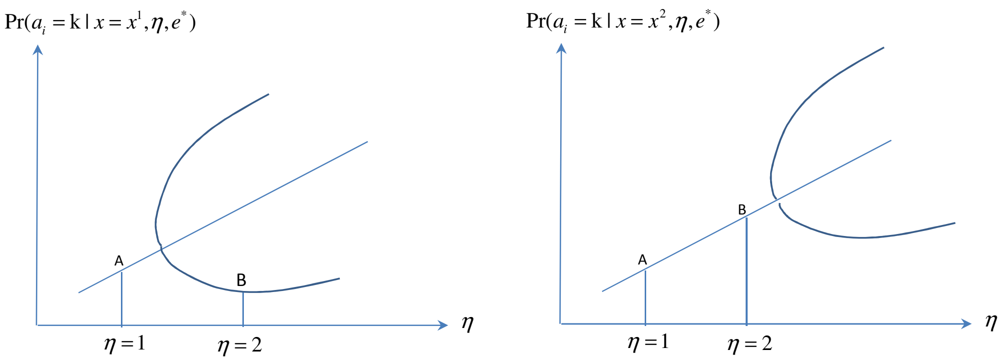

Figure B.3: The CCPs associated with different unobserved heterogeneity

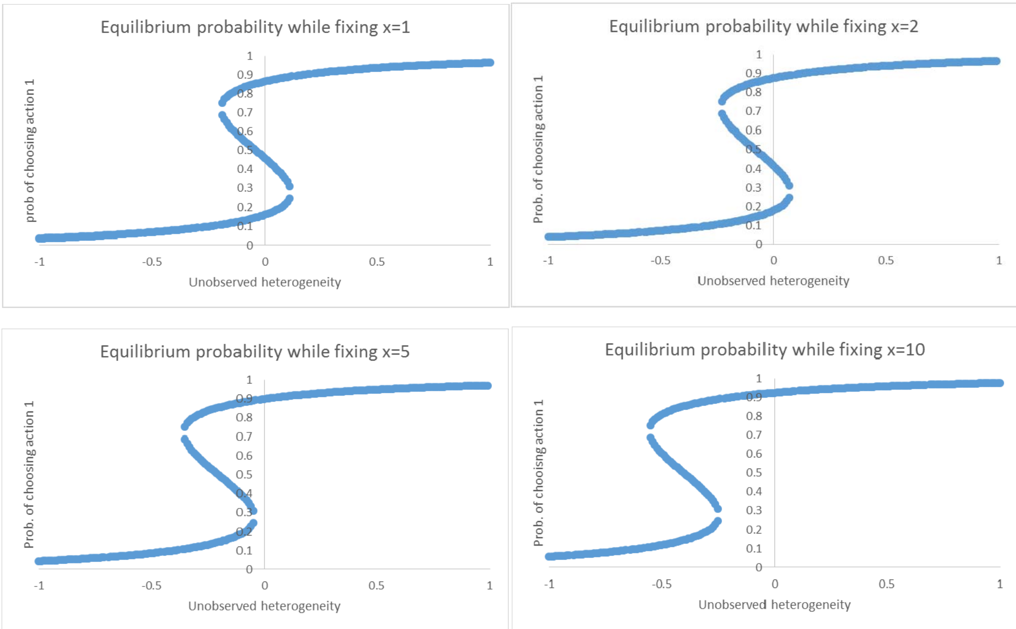

Table B.1: Estimates of CCPs for different x

	$x=1$			$x=2$			$x=3$		$x=4$	
	$\eta=0$		$\eta=1$	$\eta=0$		$\eta=0.5$	$\eta=0$	$\eta=0.5$	$\eta=0$	$\eta=0.5$
true	.159	.460	.939	.178	0.417	.942	.205	.945	.89	.948
$m=800$	.167	.571	.962	.192	0.559	.956	.199	.945	.730	.940
	(.114)	(.168)	(.044)	(.101)	(.186)	(.034)	(.063)	(.010)	(.287)	(.161)
$m=1000$	.162	.560	.963	.187	.560	.962	.206	.945	.760	.941
	(.101)	.156	(.032)	(.097)	(.194)	(.029)	(.059)	(.009)	(.285)	(.181)
$m=1500$	.162	.547	.955	0.175	.525	.958	.202	.945	.768	.943
	(.089)	(.108)	(.029)	(.093)	(.179)	(.031)	(.042)	(.007)	(.264)	(.157)
$m=3000$	.147	.491	.949	.177	.502	.953	.207	.946	.806	.955
	(.062)	(.086)	(.022)	(.075)	(.153)	(.024)	(.035)	(.005)	(.191)	(.097)
$m=5000$	.160	.487	.945	.172	.460	.946	.207	.946	.812	.961
	(.061)	(.070)	(.017)	(.064)	(.104)	(.017)	(.030)	(.004)	(.187)	(.035)

[1] The number in brackets is the standard deviation computed through bootstrap with 500 repetition
[2] For markets with unique equilibrium, δ is assumed to be zero

Appendix C

Appendix to Chapter 4

A Proofs

Proof of Lemma 4.3 $\{w_t, \eta_t\}$ follows a first-order Markov Chain.

$$\begin{aligned}
& \Pr(w_t, \eta_t | w_{t-1}, \eta_{t-1}, \Omega_{<t-1}) \\
= \; & \Pr(a_t, x_t, \eta_t | a_{t-1}, x_{t-1}, \eta_{t-1}, \Omega_{<t-1}) \\
= \; & \Pr(a_t | x_t, \eta_t, a_{t-1}, \Omega_{<t-1}) \Pr(x_t | \eta_t, x_{t-1}, \eta_{t-1}, a_{t-1}, \Omega_{<t-1}) \Pr(\eta_t | x_{t-1}, \eta_{t-1}, a_{t-1}, \Omega_{<t-1}) \\
= \; & \Pr(a_t | x_t, \eta_t, a_{t-1}) \Pr(x_t | \eta_t, x_{t-1}, a_{t-1}) \Pr(\eta_t | x_{t-1}, \eta_{t-1}, a_{t-1}) \\
= \; & \Pr(a_t, x_t, \eta_t | x_{t-1}, \eta_{t-1}, a_{t-1}) \\
= \; & \Pr(w_t, \eta_t | w_{t-1}, \eta_{t-1})
\end{aligned}$$

The third equality holds because of assumption 1 and the Markov perfect equilibrium assumption. In a given market, I can express the joint distribution of three periods of data as follows:

$$\Pr(w_{t+2}, w_{t+1}, w_t) = \sum_{\eta_{t+1}} \Pr(w_{t+2} | w_{t+1}, \eta_{t+1}) \Pr(w_{t+1}, w_t | \eta_{t+1}) \Pr(\eta_{t+1}) \tag{A.1}$$

In a given market, the joint distribution of four periods of observables can be represented as follows:

$$\begin{aligned}
& \Pr(w_{t+3}, w_{t+2}, w_{t+1}, w_t) && \text{(A.2)}\\
= & \sum_{\eta_{t+2},\eta_{t+1}} \Pr(w_{t+3}, w_{t+2}, \eta_{t+2}, w_{t+1}, \eta_{t+1}, w_t)\\
= & \sum_{\eta_{t+2},\eta_{t+1}} \Pr(w_{t+3}|w_{t+2}, \eta_{t+2}) \Pr(w_{t+2}, \eta_{t+2}|w_{t+1}, \eta_{t+1}) \Pr(w_{t+1}, \eta_{t+1}, w_t)\\
= & \sum_{\eta_{t+2},\eta_{t+1}} \Pr(w_{t+3}|w_{t+2}, \eta_{t+2}) \Pr(w_{t+2}|\eta_{t+2}, w_{t+1}, \eta_{t+1}) \Pr(\eta_{t+2}|w_{t+1}, \eta_{t+1}) \Pr(w_{t+1}, \eta_{t+1}, w_t)\\
= & \sum_{\eta_{t+2},\eta_{t+1}} \Pr(w_{t+3}|w_{t+2}, \eta_{t+2}) \Pr(w_{t+2}|\eta_{t+2}, w_{t+1}) \Pr(\eta_{t+2}|w_{t+1}, \eta_{t+1}) \Pr(w_{t+1}, \eta_{t+1}, w_t)\\
= & \sum_{\eta_{t+2},\eta_{t+1}} \Pr(w_{t+3}|w_{t+2}, \eta_{t+2}) \Pr(w_{t+2}|\eta_{t+2}, w_{t+1}) \Pr(\eta_{t+2}, w_{t+1}, \eta_{t+1}, w_t)\\
= & \sum_{\eta_{t+2}} \Pr(w_{t+3}|w_{t+2}, \eta_{t+2}) \Pr(w_{t+2}|\eta_{t+2}, w_{t+1}) \sum_{\eta_{t+1}} \Pr(\eta_{t+2}, w_{t+1}, \eta_{t+1}, w_t)\\
= & \sum_{\eta_{t+2}} \Pr(w_{t+3}|w_{t+2}, \eta_{t+2}) \Pr(w_{t+2}|\eta_{t+2}, w_{t+1}) \Pr(w_{t+1}, w_t, \eta_{t+2}) && \text{(A.3)}
\end{aligned}$$

□

Proof of Lemma 4.6 With the cardinality of unobserved factor $Q \times L$ identified, I partition the space of w_{t+3} and w_t into $Q \times L$. Fixing w_{t+2} and w_{t+1}, matrix $F_{w_{t+3},w_{t+2},w_{t+1},w_t}$ defined as below constructed is invertible. As a result, matrices $A_{w_{t+3}|w_{t+2},\tau_{t+2}}$ and $B_{w_{t+1},w_t,\tau_{t+2}}$ 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,

$$\begin{aligned}
(w_{t+2}, w_{t+1}) &: F_{w_{t+3},w_{t+2},w_{t+1},w_t} = A_{w_{t+3}|w_{t+2},\tau_{t+2}} D_{w_{t+2}|w_{t+1},\tau_{t+2}} B_{w_{t+1},w_t,\tau_{t+2}} && \text{(A.4)}\\
(\bar{w}_{t+2}, w_{t+1}) &: F_{w_{t+3},\bar{w}_{t+2},w_{t+1},w_t} = A_{w_{t+3}|\bar{w}_{t+2},\tau_{t+2}} D_{\bar{w}_{t+2}|w_{t+1},\tau_{t+2}} B_{w_{t+1},w_t,\tau_{t+2}} && \text{(A.5)}\\
(w_{t+2}, \bar{w}_{t+1}) &: F_{w_{t+3},w_{t+2},\bar{w}_{t+1},w_t} = A_{w_{t+3}|w_{t+2},\tau_{t+2}} D_{w_{t+2}|\bar{w}_{t+1},\tau_{t+2}} B_{\bar{w}_{t+1},w_t,\tau_{t+2}} && \text{(A.6)}\\
(\bar{w}_{t+2}, \bar{w}_{t+1}) &: F_{w_{t+3},\bar{w}_{t+2},\bar{w}_{t+1},w_t} = A_{w_{t+3}|\bar{w}_{t+2},\tau_{t+2}} D_{\bar{w}_{t+2}|\bar{w}_{t+1},\tau_{t+2}} B_{\bar{w}_{t+1},w_t,\tau_{t+2}} && \text{(A.7)}
\end{aligned}$$

Matrices $A_{w_{t+3}|w_{t+2},\tau_{t+2}}$ and $B_{w_{t+1},w_t,\tau_{t+2}}$ are invertible by construction. Assume that $Pr(w_{t+2}|w_{t+1}, \tau_{t+2})$ is positive for every combination of w_{t+2} and w_{t+1}, so matrix $D_{w_{t+2}|w_{t+1},\tau_{t+2}}$ is also invertible. Consequently, we can post-multiply inverse of equation A.5 to equation A.4, to obtain:

$$Y \equiv F_{w_{t+3},w_{t+2},w_{t+1},w_t} F^{-1}_{w_{t+3},\bar{w}_{t+2},w_{t+1},w_t} = A_{w_{t+3}|w_{t+2},\tau_{t+2}} D_{w_{t+2}|w_{t+1},\tau_{t+2}} D^{-1}_{\bar{w}_{t+2}|w_{t+1},\tau_{t+2}} A^{-1}_{w_{t+3}|\bar{w}_{t+2},\tau_{t+2}} \quad \text{(A.8)}$$

Similarly,

$$Z \equiv F_{w_{t+3},\bar{w}_{t+2},\bar{w}_{t+1},w_t} F^{-1}_{w_{t+3},w_{t+2},\bar{w}_{t+1},w_t} = A_{w_{t+3}|\bar{w}_{t+2},\tau_{t+2}} D_{\bar{w}_{t+2}|\bar{w}_{t+1},\tau_{t+2}} D^{-1}_{w_{t+2}|\bar{w}_{t+1},\tau_{t+2}} A^{-1}_{w_{t+3}|w_{t+2},\tau_{t+2}} \quad \text{(A.9)}$$

Consequently, I postmultiply Eq. A.8 by Eq. A.9, leading to

$$\begin{aligned} YZ &= A_{w_{t+3}|w_{t+2},\tau_{t+2}} \left(D_{w_{t+2}|w_{t+1},\tau_{t+2}} D^{-1}_{\bar{w}_{t+2}|w_{t+1},\tau_{t+2}} D_{\bar{w}_{t+2}|\bar{w}_{t+1},\tau_{t+2}} D^{-1}_{w_{t+2}|\bar{w}_{t+1},\tau_{t+2}} \right) A^{-1}_{w_{t+3}|w_{t+2},\tau_{t+2}} \\ &\equiv A_{w_{t+3}|w_{t+2},\tau_{t+2}} D_{w_{t+2},\bar{w}_{t+2},w_{t+1},\bar{w}_{t+1}|\tau_{t+2}} A^{-1}_{w_{t+3}|w_{t+2},\tau_{t+2}} \qquad where \qquad (A.10) \end{aligned}$$

$$\begin{aligned} D_{w_{t+2},\bar{w}_{t+2},w_{t+1},\bar{w}_{t+1}|\tau_{t+2}} &= D_{w_{t+2}|w_{t+1},\tau_{t+2}} D^{-1}_{\bar{w}_{t+2}|w_{t+1},\tau_{t+2}} D_{\bar{w}_{t+2}|\bar{w}_{t+1},\tau_{t+2}} D^{-1}_{w_{t+2}|\bar{w}_{t+1},\tau_{t+2}} \\ &= \frac{Pr(w_{t+2}|w_{t+1},\tau_{t+2})Pr(\bar{w}_{t+2}|\bar{w}_{t+1},\tau_{t+2})}{Pr(\bar{w}_{t+2}|w_{t+1},\tau_{t+2})Pr(w_{t+2}|\bar{w}_{t+1},\tau_{t+2})} \\ &\equiv C(w_{t+2},\bar{w}_{t+2},w_{t+1},\bar{w}_{t+1}|\tau_{t+2}) \end{aligned}$$

Equation A.10 results in a eigenvalue-eigenvector decomposition for observed matrix YZ, with eigenvectors corresponding to the matrix $A_{w_{t+3}|w_{t+2},\tau_{t+2}}$ and the eigenvalues corresponding to matrix $D_{w_{t+2},\bar{w}_{t+2},w_{t+1},\bar{w}_{t+1}|\tau_{t+2}}$.

Evaluating w_{t+2} to another value $\bar{w}_{t+2}$, I can obtain another matrix expression that share the common term $B_{w_{t+1},w_t,\tau_{t+2}}$ with equation 4.3. By the same logic, changing the value of $\bar{w}_{t+1}$, evaluating w_{t+2} in two different values results in two matrix equations sharing the common term $B_{\bar{w}_{t+1},w_t,\tau_{t+2}}$. Using this feature, manipulation over these four matrix expression leads to a matrix eigen-decomposition expression. □

Proof of Lemma 4.8: Identification of Law of Motion Again, with four periods of data, the joint distribution of observables can be expressed to be factorized as the components that we want to identify in the followings:

$$\begin{aligned} \Pr(w_{t+3},w_{t+2},w_{t+1},w_t) &= \sum_{\tau_{t+2}} \Pr(w_{t+3}|w_{t+2},\tau_{t+2}) \Pr(w_{t+2},\tau_{t+2},w_{t+1},w_t) \quad (A.11) \\ \Pr(w_{t+2},\tau_{t+2},w_{t+1},w_t) &= \sum_{\tau_{t+1}} \Pr(w_{t+2},\tau_{t+2},w_{t+1},\tau_{t+1},w_t) \\ &= \sum_{\tau_{t+1}} \Pr(w_{t+2},\tau_{t+2}|w_{t+1},\tau_{t+1}) \Pr(w_{t+1},\tau_{t+1},w_t) \quad (A.12) \end{aligned}$$

Fixed $w_{t+2} = \bar{w}_{t+2}$ and $w_{t+1} = \bar{w}_{t+1}$ and rewrite above equations into similar matrix format defined at the beginning:

$$F_{w_{t+3},w_{t+2},w_{t+1},w_t} = A_{w_{t+3}|\bar{w}_{t+2},\tau_{t+2}} B_{\bar{w}_{t+2},\tau_{t+2},\bar{w}_{t+1},w_t} \quad (A.13)$$

$$B_{\bar{w}_{t+2},\tau_{t+2},\bar{w}_{t+1},w_t} = A_{\bar{w}_{t+2},\tau_{t+2}|\bar{w}_{t+1},\tau_{t+1}} A_{\bar{w}_{t+1},\tau_{t+1},w_t} \quad (A.14)$$

Consequently, we have:

$$F_{w_{t+3},w_{t+2},w_{t+1},w_t} = A_{w_{t+3}|\bar{w}_{t+2},\tau_{t+2}} A_{\bar{w}_{t+2},\tau_{t+2}|\bar{w}_{t+1},\tau_{t+1}} A_{\bar{w}_{t+1},\tau_{t+1},w_t} \quad (A.15)$$

I show in lemma 4.6 that $A_{w_{t+3}|\bar{w}_{t+2},\tau_{t+2}}$ is identified. Additionally, the left-land side matrix can be computed from the data. If $A_{\bar{w}_{t+1},\tau_{t+1},w_t}$ is identified, the law of motion is identified given both $A_{\bar{w}_{t+1},\tau_{t+1},w_t}$ and $A_{w_{t+3}|\bar{w}_{t+2},\tau_{t+2}}$ are invertible by the way I construct the two matrices. I next show that $A_{\bar{w}_{t+1},\tau_{t+1},w_t}$ can be identified through following equation.

$$\Pr(w_{t+2}, w_{t+1}, w_t) = \sum_{\tau_{t+1}} \Pr(w_{t+2}|w_{t+1}, \tau_{t+1}) \Pr(w_{t+1}, \tau_{t+1}, w_t) \tag{A.16}$$

Similar logic, fixing $w_{t+1} = \bar{w}_{t+1}$, above equation's matrix counterpart is as follows:

$$F_{w_{t+2},\bar{w}_{t+1},w_t} = A_{w_{t+2}|\bar{w}_{t+1},\tau_{t+1}} A_{\bar{w}_{t+1},\tau_{t+1},w_t} \tag{A.17}$$

Since $A_{w_{t+2}|\bar{w}_{t+1},\tau_{t+1}}$ is identified and invertible, $A_{\bar{w}_{t+1},\tau_{t+1},w_t}$ is identified. Consequently, the law of motion is identified. □

Proof of lemma 4.9: Identification of Initial Condition Given that we already identified the transition matrix $Pr(w_{t+1}|w_t, \tau_t)$, following equation provides identification of the initial distribution:

$$\Pr(w_{t+1}, w_t) = \sum_{\tau_t} \Pr(w_{t+1}|w_t, \tau_t) \Pr(w_t, \tau_t) \tag{A.18}$$

fixing $w_t = \bar{w}_t$ and rewrite above equation in the matrix format:

$$V_{w_{t+1},\bar{w}_t} = A_{w_{t+1}|\bar{w}_t,\tau_t} V_{\bar{w}_t,\tau_t} \tag{A.19}$$

Since $A_{w_{t+1}|\bar{w}_t,\tau_t}$ is invertible, $\Pr(w_t, \tau_t)$ is identified for all w_t. □

Proof of lemma 4.12: Identification of payoff functions Note that the equilibrium condition with extreme value distribution assumption on the private shocks becomes:

$$log(\sigma_i(a_i = k, s)) - log(\sigma_i(a_i = 0, s)) = V_i(a_i = k, s) - V_i(a_i = 0, s) \tag{A.20}$$

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

$$\begin{aligned} V_i(s) &= E_{\epsilon_i} max_{a_i} V_i(a_i, s) + \epsilon_i(a_i) = log(\sum_{k=0}^{K} exp(V_i(k, s))) \\ &= log \sum_{k=0}^{K} exp(V_i(k, s) - V_i(0, s)) + V_i(0, s) \end{aligned} \tag{A.21}$$

The second equation holds because of i.i.d. and the extreme value distribution assumption of ϵ_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:

$$\begin{aligned} V_i(a_i, s) &= \Pi_i(a_i, s) + \beta E(V_i(s')|s, a_i) \\ &= \Pi_i(a_i, s) + \beta E(log \sum_{k=0}^{K} exp(V_i(k, s) - V_i(0, s)) + V_i(0, s')|s, a_i) \\ &= \Pi_i(a_i, s) + \beta E(log \sum_{k=0}^{K} exp(V_i(k, s) - V_i(0, s))) + \beta E(V_i(0, s')|s, a_i) \end{aligned} \tag{A.22}$$

With the normalization assumption, for the action that I imposed the zero period utility, I have:

$$\begin{aligned} V_i(0,s) &= \Pi_i(a_i=0,s) + \beta E(log\sum_{k=0}^{K} exp(V_i(k,s)-V_i(0,s))) + \beta E(V_i(0,s')|s,a_i=0) \\ &= \beta E(log\sum_{k=0}^{K} exp(V_i(k,s)-V_i(0,s))) + \beta E(V_i(0,s')|s,a_i=0) \end{aligned} \quad (A.23)$$

By the equilibrium condition, the term $(log\sum_{k=0}^{K} exp(V_i(k,s)-V_i(0,s)))$ can be computed through $log(\sigma_i(a_i=k,s))-log(\sigma_i(a_i=0,s))$, so we can treat it as a constant. Consequently, equation A.23 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 A.21, the ex ante value function $V_i(s)$ is identified. From equation A.23, $V_i(a_i=k,s)$ is identified. Then we can identify the expected period payoff $\Pi_i(a_i,s)$ through equation A.22 as:

$$\Pi_i(a_i,s) = V_i(a_i,s) - \beta E(log\sum_{k=0}^{K} exp(V_i(k,s)-V_i(0,s))) - \beta E(V_i(0,s')|s,a_i) \quad (A.24)$$

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 restrictions. $\square$

B Graphs and Tables

Figure C.1: The Estimation of the Number of Equilibria

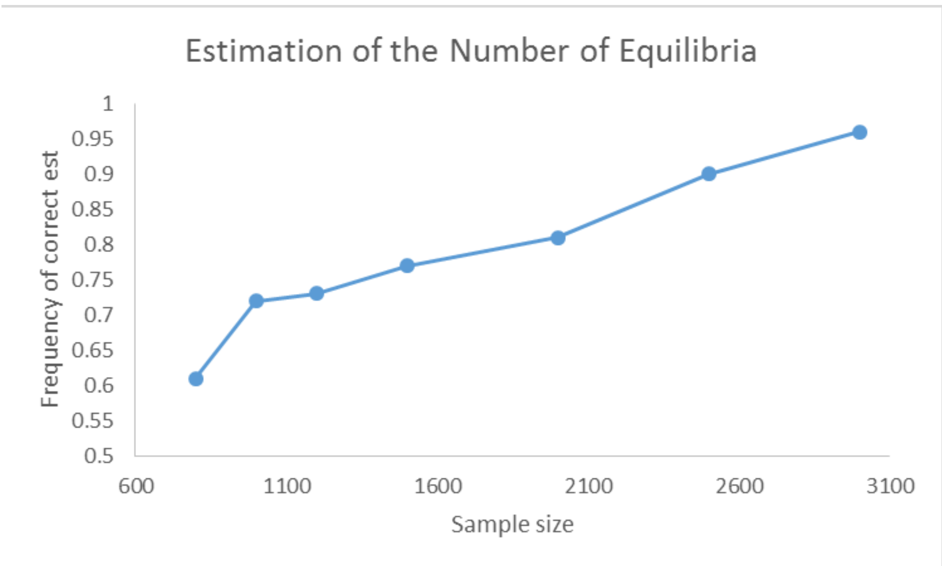

Table C.1: Estimates of Law of Motion

	True Values		Ests (ns=1000)		Ests (ns=2000)	
$\Pr(w^*_{t+1} \mid w_t = 1, e^*)$	0.72	0.47	0.6728	0.3865	0.6906	0.4186
			(0.2147)	(0.2064)	(0.2145)	(0.2196)
	0.28	0.53	0.3272	0.6135	0.3094	0.5814
			(0.2147)	(0.2064)	(0.2145)	(0.2196)
$\Pr(w^*_{t+1} \mid w_t = 2, e^*)$	0.78	0.70	0.7723	0.5032	0.7756	0.6263
			(0.1763)	(0.3307)	(0.1303)	(0.2701)
	0.22	0.30	0.2277	0.4968	0.2244	0.3737
			(0.1763)	(0.3307)	(0.1303)	(0.2701)
$\Pr(w^*_{t+1} \mid w_t = 3, e^*)$	0.58	0.16	0.5936	0.2876	0.5711	0.2503
			(0.2431)	(0.2135)	(0.2460)	(0.2050)
	0.42	0.84	0.4064	0.7124	0.4289	0.7498
			(0.2431)	(0.2135)	(0.2460)	(0.2050)
$\Pr(w^*_{t+1} \mid w_t = 4, e^*)$	0.71	0.42	0.5939	0.3019	0.6713	0.3538
			(0.2736)	(0.2302)	(0.2523)	(0.2285)
	0.29	0.58	0.4061	0.6981	0.3287	0.6462
			(0.2736)	(0.2302)	(0.2523)	(0.2285)

[1] The number in brackets is the standard deviation computed through bootstrap with replacement, with 100 repetition

Bibliography

Ackerberg, D. A., and G. Gowrisankaran (2006): "Quantifying equilibrium network externalities in the ACH banking industry," *The RAND journal of economics*, 37(3), 738–761.

Aguirregabiria, V., and P. Mira (2002): "Identification and Estimation of Dynamics Discrete Games," in *manuscript presented at the Conference on Industrial Organization and the Food Processing Industry, Institut dEconomie Industrielle (IDEI). Toulouse, France.*

——— (2007): "Sequential estimation of dynamic discrete games," *Econometrica*, 75(1), 1–53.

——— (2013): "Identification of Games of Incomplete Information with Multiple Equilibria and Common Unobserved Heterogeneity," Discussion paper.

Ai, C., and X. Chen (2003): "Efficient estimation of models with conditional moment restrictions containing unknown functions," *Econometrica*, 71(6), 1795–1843.

Andrew, A. L., K.-W. E. Chu, and P. Lancaster (1993): "Derivatives of eigenvalues and eigenvectors of matrix functions," *SIAM journal on matrix analysis and applications*, 14(4), 903–926.

ANDREWS, D. W., S. BERRY, AND P. JIA (2004): "Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store location," .

ARADILLAS-LOPEZ, A. (2010): "Semiparametric estimation of a simultaneous game with incomplete information," *Journal of Econometrics*, 157(2), 409–431.

ARCIDIACONO, P., AND R. A. MILLER (2011): "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity," *Econometrica*, 79(6), 1823–1867.

BAJARI, P., C. L. BENKARD, AND J. LEVIN (2007): "Estimating dynamic models of imperfect competition," *Econometrica*, 75(5), 1331–1370.

BAJARI, P., V. CHERNOZHUKOV, H. HONG, AND D. NEKIPELOV (2009): "Identification and efficient semiparametric estimation of a dynamic discrete game," Discussion paper, Working paper, University of Minnesota.[203].

BAJARI, P., J. HAHN, H. HONG, AND G. RIDDER (2011): "A note on semiparametric estimation of finite mixtures of discrete choice models with application to game theoretic models*," *International Economic Review*, 52(3), 807–824.

BAJARI, P., H. HONG, J. KRAINER, AND D. NEKIPELOV (2010a): "Computing equilibria in static games of incomplete information using the all-solution homotopy," *Operations Research*, 58.

——— (2010b): "Estimating static models of strategic interactions," *Journal of Business & Economic Statistics*, 28(4).

Bajari, P., H. Hong, and S. Ryan (2010): "Identification and estimation of a discrete game of complete information," *Econometrica*, 78(5), 1529–1568.

Beresteanu, A., P. Ellickson, and S. Misra (2010): "The dynamics of retail oligopoly," *University of Pittsburgh and University of Rochester.*

Berk, R., L. Brown, and L. Zhao (2010): "Statistical inference after model selection," *Journal of Quantitative Criminology*, 26(2), 217–236.

Berry, S. (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica: Journal of the Econometric Society*, pp. 889–917.

Berry, S., and P. Reiss (2007): "Empirical models of entry and market structure," *Handbook of industrial organization*, 3, 1845–1886.

Berry, S., and E. Tamer (2006): "Identification in models of oligopoly entry," *ECONOMETRIC SOCIETY MONOGRAPHS*, 42, 46.

Bjorn, P. A., and Q. H. Vuong (1984): "Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation," Discussion paper.

Bresnahan, T. F., and P. C. Reiss (1990): "Entry in monopoly market," *The Review of Economic Studies*, 57(4), 531–553.

——— (1991): "Empirical models of discrete games," *Journal of Econometrics*, 48(1), 57–81.

Camba-Mendez, G., and G. Kapetanios (2009): "Statistical tests and estimators of the rank of a matrix and their applications in econometric modelling," *Econometric Reviews*, 28(6), 581–611.

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

De Paula, A. (2012): "Econometric analysis of games with multiple equilibria," *Annual Review of Economics*, (0).

De Paula, A., and X. Tang (2012): "Inference of signs of interaction effects in simultaneous games with incomplete information," *Econometrica*, 80(1), 143–172.

Echenique, F., and I. Komunjer (2009): "Testing models with multiple equilibria by quantile methods," *Econometrica*, 77(4), 1281–1297.

Egesdal, M., Z. Lai, and C.-L. Su (2013): "Estimating Dynamic Discrete-Choice Games of Incomplete Information," *Available at SSRN 2157329.*

Ellickson, P. B., and S. Misra (2008): "Supermarket pricing strategies," *Marketing Science*, 27(5), 811–828.

Ericson, R., and A. Pakes (1995): "Markov-perfect industry dynamics: A framework for empirical work," *The Review of Economic Studies*, 62(1), 53–82.

Grieco, P. L. (2011): "Discrete Games with Flexible Information Structures: An Application to Local Grocery Markets," *Available at SSRN 1859643.*

Haller, H., and R. Lagunoff (2000): "Genericity and Markovian behavior in stochastic games," *Econometrica*, 68(5), 1231–1248.

Harsanyi, J. C. (1973): "Oddness of the number of equilibrium points: a new proof," *International Journal of Game Theory*, 2(1), 235–250.

Holmes, T. J. (2011): "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 79(1), 253–302.

Hosoya, Y. (1989): "Hierarchical Statistical Models and a Generalized Likelihood Ratio Test," *Journal of the Royal Statistical Society. Series B (Methodological)*, 51(3), pp. 435–447.

Hotz, V. J., and R. A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models," *The Review of Economic Studies*, 60(3), 497–529.

Hu, Y. (2008): "Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution," *Journal of Econometrics*, 144(1), 27–61.

Hu, Y., D. McAdams, and M. Shum (2013): "Identification of first-price auctions with non-separable unobserved heterogeneity," *Journal of Econometrics*.

Hu, Y., and M. Shum (2012): "Nonparametric identification of dynamic models with unobserved state variables," *Journal of Econometrics*, 171(1), 32–44.

Jia, P. (2008): "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 76(6), 1263–1316.

Jovanovic, B. (1989): "Observable implications of models with multiple equilibria," *Econometrica: Journal of the Econometric Society*, pp. 1431–1437.

KASAHARA, H., AND K. SHIMOTSU (2009): "Nonparametric identification of finite mixture models of dynamic discrete choices," *Econometrica*, 77(1), 135–175.

——— (2012): "Sequential estimation of structural models with a fixed point constraint," *Econometrica*, 80(5), 2303–2319.

KHAN, S., AND E. TAMER (2010): "Irregular identification, support conditions, and inverse weight estimation," *Econometrica*, 78(6), 2021–2042.

KLEIBERGEN, F., AND R. PAAP (2006): "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 133(1), 97–126.

MATZKIN, R. L. (1992): "Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models," *Econometrica: Journal of the Econometric Society*, pp. 239–270.

MCDOWELL, W., AND S. J. DICK (2003): "Switching radio stations while driving: Magnitude, motivation, and measurement issues," *Journal of Radio Studies*, 10(1), 46–62.

MILGROM, P. R., AND R. J. WEBER (1985): "Distributional strategies for games with incomplete information," *Mathematics of Operations Research*, 10(4), 619–632.

MULLER, H.-G. (1992): "Change-points in nonparametric regression analysis," *The Annals of Statistics*, 20(2), 737–761.

NEVO, A., AND V. AGUIRREGABIRIA (2010): "Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games," *Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games (December 29, 2010). Northwestern University Center for the Study of Industrial Organization Working Paper*, (0107).

Newey, W. K. (1990): "Semiparametric efficiency bounds," *Journal of applied econometrics*, 5(2), 99–135.

——— (1994): "The asymptotic variance of semiparametric estimators," *Econometrica: Journal of the Econometric Society*, pp. 1349–1382.

Pakes, A., M. Ostrovsky, and S. Berry (2007): "Simple estimators for the parameters of discrete dynamic games (with entry/exit examples)," *The RAND Journal of Economics*, 38(2), 373–399.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2006): "Moment inequalities and their application," *Unpublished Manuscript*.

Pesendorfer, M., and P. Schmidt-Dengler (2008): "Asymptotic least squares estimators for dynamic games," *The Review of Economic Studies*, 75(3), 901–928.

Pesendorfer, M., and Y. Takahashi (2012): "Testing for Equilibrium Multiplicity in Dynamic Markov Games," .

Robin, J.-M., and R. J. Smith (2000): "Tests of rank," *Econometric Theory*, 16(02), 151–175.

Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica: Journal of the Econometric Society*, pp. 999–1033.

Ryan, S. P. (2012): "The costs of environmental regulation in a concentrated industry," *Econometrica*, 80(3), 1019–1061.

Seim, K. (2006): "An empirical model of firm entry with endogenous product-type choices," *The RAND Journal of Economics*, 37(3), 619–640.

Su, C.-L. (2012): "Estimating Discrete-Choice Games of Incomplete Information: A Simple Static Example," *Available at SSRN 1997548.*

Sweeting, A. (2009): "The strategic timing incentives of commercial radio stations: An empirical analysis using multiple equilibria," *The RAND Journal of Economics*, 40(4), 710–742.

——— (2011): "Dynamic product positioning in differentiated product markets: the effect of fees for musical performance rights on the commercial radio industry," *Manuscript. Duke University.*

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

Curriculum Vitae

Ruli Xiao was born in a small town in the south of China in 1984. She received her B.S. in Statistics at Tongji University, Shanghai, China in 2006. Ruli entered the School of Economics at Shanghai University of Finance and Economics and earned her M.A. 2008. Ruli started her study in the Ph.D. program of economics at Johns Hopkins University from the fall of 2008. She will start as an assistant professor in the department of economics at Indiana University from August 2014.